COURS DE L'ÉCOLE D'ÉTÉ DE PHYSIQUE THÉORIQUE

QUANTUM MECHANICS

Bryce S. DE WITT

Radiation Laboratory
University of California

★

LES HOUCHES
ÉTÉ 1953

Q U A N T U M M E C H A N I C S

Lecture Notes by

Bryce S. DE WITT
Radiation Laboratory
University of California

Chapters I & II

Université de Grenoble

Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

1. A.- REVIEW OF CLASSICAL MECHANICS.

a) Canonical transformations and invariants.

We consider a system having n coordinates q^i (n may be non-denumerably infinite !) satisfying equations of motion of the form

$$0 = \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \tag{1.1}$$

where L is an explicit function (the Lagrangian function) of the coordinates q^i, their time derivatives $\dot{q}^i$, and perhaps also of the time t. We introduce the conjugate momenta

$$p_i = \frac{\partial L}{\partial \dot{q}^i} \tag{1.2}$$

and the Hamiltonian function

$$H = p_i \dot{q}^i - L \,. \tag{1.3}$$

We assume that equations (1.2) can be solved for the $\dot{q}^i$ in terms of the p_i, q^i and t. (This means that L must be at least quadratic in the $\dot{q}^i$.) The Hamiltonian function may then be expressed as an explicit function of the q^i, p_i and t, and we have

$$H = \dot{q}^i \delta p_i + p_i \delta \dot{q}^i - \frac{\partial L}{\partial q^i}\delta q^i - \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i = \dot{q}^i \delta p_i - \dot{p}_i \delta q^i \tag{1.4}$$

where the symbol δ denotes a variation of the q^i (and hence also of $\dot{q}^i$ and p_i) from an actual trajectory satisfying (1.1). It is evident that the equations of motion may be rewritten in the following form.

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \quad , \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{1.5}$$

The form of the equations (1.5) is said to be the canonical form of the equations of motion.

The Lagrangian function L may evidently be reexpressed as an explicit function $\bar{L}$ of the q^i, p_i, $\dot{q}^i$, $\dot{p}_i$ and t :

$$\bar{L}(q^i, p_i, \dot{q}^i, \dot{p}_i, t) = p_i \dot{q}^i - H(q^i, p_i, t) \qquad (1.6)$$

(L is actually independent of the $\dot{p}_i$, i.e. $\frac{\partial \bar{L}}{\partial \dot{p}_i} = 0$) In terms of the function L the equations of motion may be written in the form

$$\left.\begin{aligned} 0 &= \frac{\partial \bar{L}}{\partial q^i} - \frac{d}{dt}\frac{\partial \bar{L}}{\partial \dot{q}^i} \\ 0 &= -\frac{\partial \bar{L}}{\partial p_i} - \frac{d}{dt}\frac{\partial \bar{L}}{\partial \dot{p}_i} \end{aligned}\right\} \qquad (1.7)$$

Problem I : Verify equations (1.7)

Equations (1.7) have again the Lagrangian form and result from the varaitional principle

$$\delta \int \bar{L}\, dt = 0 \qquad (1.8)$$

just as the equations (1.1) result from the variational principle

$$\delta \int L\, dt = 0 . \qquad (1.9)$$

Now, the form of the Lagrangian equations (1.7) remains invariant under any transformation of the q^i and p_i of the general form

$$\left.\begin{aligned} q'^i &= q'^i(q, p, t) \\ p'_i &= p'_i(q, p, t) \end{aligned}\right\} \qquad (1.10)$$

with non-vanishing Jacobian

$$\frac{\partial(p', q')}{\partial(p, q)} \neq 0, \tag{1.11}$$

just as the form of the equations (1.1) remains invariant under transformations of the form

$$p'^i = p'^i(p, q, t) \tag{1.12}$$

However, we shall be interested only in those transformations (1.10) which leave invariant not only the form of equations (1.7) but also the canonical form of the equations (1.5). Such transformations are evidently* those for which

$$p_i \dot{q}^i - H(p, q, t) = p'_i \dot{q}'^i - H'(p', q', t) + \dot{V} \tag{1.13}$$

where H' is some function of the new variables (and perhaps also of the time) and V is a function of any independent $2n$ of the $4n$ variables q^i, p_i, q'^i, p'_i (and perhaps also of the time). For, if we write

$$\bar{L}'(p', q', \dot{p}', \dot{q}', t) = p'_i \dot{q}'^i - H'(p', q', t), \tag{1.14}$$

we have

$$\delta \int \bar{L}' dt = \delta \int \bar{L} dt \tag{1.15}$$

since $\bar{L}'$ and $\bar{L}$ differ by a total time derivative $\dot{V}$. The new equations

---

* It should be stated that for any given system there may be other transformations which will lead to new equations which can also be put into canonical form. Such transformations will, however, be special to the particular system under consideration. We consider here only those transformations which leave the canonical form of all Hamiltonian systems of equations simultaneously invariant.

of motion are

$$\dot{q}'^i = \frac{\partial H'}{\partial p'_i} \quad , \quad \dot{p}'_i = -\frac{\partial H}{\partial q'^i} \tag{1.16}$$

H' is the new Hamiltonian function.

A transformation (1.10) which leaves invariant the canonical form of the equations (1.5) is called a canonical tranformation. A special type of canonical transformation is given by the transformation (1.12) together with

$$p'_i = \frac{\partial q^j}{\partial q'^i} p_j \tag{1.17}$$

(In this case $H'(q', p', t) = H(q, p, t) + p'_i \frac{\partial q'^i}{\partial t}$)

Equations (1.12) and (1.17) together define what is known as a point transformation. The set of all point transformations forms a subgroup of the group of all canonical transformations, which in turn is a subgroup of the group of all transformations (1.10).

We now discuss four principal forms for canonical transformations :

I) $V = \Omega_1 (q, q', t)$

In this case

$$0 = p_i \dot{q}^i - H - p'_i \dot{q}'^i + H' - \frac{\partial \Omega_1}{\partial q^i} \dot{q}^i - \frac{\partial \Omega_1}{\partial q'^i} \dot{q}'^i - \frac{\partial \Omega_1}{\partial t}$$

and we have

$$p_i = \frac{\partial \Omega_1}{\partial q^i} \quad , \quad p'_i = -\frac{\partial \Omega_1}{\partial q'^i} \quad , \quad H' = H + \frac{\partial \Omega_1}{\partial t} \tag{1.18}$$

II) $V = \Omega_2 (q, p', t) - p'_i q'^i$

In this case

$$0 = p_i \dot{q}^i - H - p'_i \dot{q}'^i + H' + p'_i \dot{q}'^i + q'^i \dot{p}'_i - \frac{\partial \Omega_2}{\partial q^i} \dot{q}^i - \frac{\partial \Omega_2}{\partial p'_i} \dot{p}'_i - \frac{\partial \Omega_2}{\partial t}$$

and we have

$$p_i = \frac{\partial \Omega_2}{\partial q^i} \quad , \quad q'^i = \frac{\partial \Omega_2}{\partial p'_i} \quad , \quad H' = H + \frac{\partial \Omega_2}{\partial t} \tag{1.19}$$

III) $V = \Omega_3 (p, p', t) + p_i q'^i$

In this case

$$0 = p_i \dot{q}^i - H - p'_i \dot{q}'^i + H' - p_i \dot{q}^i - p_i q'^i - \frac{\partial \Omega_3}{\partial p_i} \dot{p}_i - \frac{\partial \Omega_3}{\partial q'^i} \dot{q}'^i - \frac{\partial \Omega_3}{\partial t}$$

and we have

$$q^i = -\frac{\partial \Omega_3}{\partial p_i}, \quad p'_i = -\frac{\partial \Omega_3}{\partial q'^i}, \quad H' = H + \frac{\partial \Omega_3}{\partial t}. \qquad (1.20)$$

IV) $V = \Omega_4 (p, p', t) - p'_i q'^i + p_i q^i$

In this case

$$0 = p_i \dot{q}^i - H - p'_i \dot{q}'^i + H' + p'_i \dot{q}'^i + p'_i \dot{q}'^i - p_i \dot{q}^i - p_i \dot{q}^i - \frac{\partial \Omega_4}{\partial p_i} \dot{p}_i - \frac{\partial \Omega_4}{\partial p'_i} \dot{p}'_i - \frac{\partial \Omega_4}{\partial t}$$

and we have

$$q^i = -\frac{\partial \Omega_4}{\partial p_i}, \quad q'^i = \frac{\partial \Omega_4}{\partial p'_i}, \quad H' = H + \frac{\partial \Omega_4}{\partial t} \qquad (1.21)$$

The point transformation (1.12, 17) is evidently given by (1.19) with

$$\Omega_2 = p'_i q'^i (p, t)$$

Problem II : Prove that the Jacobian of a point transformation is equal to unity.

The relation (1.13) may evidently be rewritten in the form

$$p_i \delta q^i - H \delta t = p'_i \delta q'^i - H' \delta t + \delta V \qquad (1.22)$$

where δq^i, δp_i, δt are completely arbitrary variations in the q^i, p_i, t, respectively and $\delta q'^i$, $\delta p'_i$, δV are corresponding variations

in the q'^i, p'_i, V. If $\delta t = 0$, we have

$$p_i \delta q^i = p'_i \delta q'^i + \delta V \qquad (1.23)$$

If the symbol d denotes a variation independent of δ we may write

$$d p_i \delta q^i + p_i d\,\delta q^i = d p'_i \delta q'^i + p'_i d \delta q'^i + d\delta V \qquad (1.24)$$

Similarly

$$\delta p_i\, d q^i + p_i \delta d q^i = \delta p'_i\, d q'^i + p'_i \delta d q'^i + \delta d V \qquad (1.25)$$

Since δ and d are independant variations we may write

$$d \delta q^i = \delta d q^i \quad , \quad d \delta p_i = \delta d p_i \qquad (1.26)$$

in the same sense in which we write

$$\frac{\partial^2 q^i}{\partial \alpha_\mu \partial \alpha_\nu} = \frac{\partial^2 q^i}{\partial \alpha_\nu \partial \alpha_\mu} \quad , \quad \frac{\partial^2 p_i}{\partial \alpha_\mu \partial \alpha_\nu} = \frac{\partial^2 p_i}{\partial \alpha_\nu \partial \alpha_\mu} \qquad (1.27)$$

when we imagine the canonical variables to be functions of 2n parameters α_κ. Equations (1.26) imply

$$d \delta q'^i = \delta d q'^i, \quad d \delta p'_i = \delta d p'_i, \quad d\delta V = \delta d V \qquad (1.28)$$

Problem III : Verify equations (1.28)

Hence, subtracting (1.24) from (1.25), we have

$$\delta p_i\, d q^i - d p_i\, \delta q^i = \delta p'_i\, d q'^i - d p'_i\, \delta q'^i \qquad (1.29)$$

We now see that not only is the form of the canonical equations (1.5) left invariant under canonical transformations but also the form of the bilinear expression $\delta p_i\, dq^i - dp_i\, \delta q^i$. There are many other expresssions which remain invariant under canonical transformations, and we shall next proceed to obtain some of these.

Let us first write

$$\begin{aligned}\delta p_i\, dq^i - dp_i\, \delta q^i &= \left(\frac{\partial p_i}{\partial q'^j}\,\delta q'^j + \frac{\partial p_i}{\partial p'_j}\,\delta p'_j\right)\left(\frac{\partial q^i}{\partial q'^k}\,dq'^k + \frac{\partial q^i}{\partial p'_k}\,dp'_k\right)\\ &\quad - \left(\frac{\partial p_i}{\partial q'^k}\,dq'^k + \frac{\partial p_i}{\partial p'_k}\,dp'_k\right)\left(\frac{\partial q^i}{\partial q'^j}\,\delta q'^j + \frac{\partial q^i}{\partial p'_j}\,\delta p'_j\right)\\ &= -((q'^j, q'^k))_{q,p}\,\delta q'^j\, dq'^k - ((q'^j, p'_k))_{q,p}\,\delta q'^j\, dp'_k\\ &\quad + ((q'^k, p'_j))_{q,p}\,\delta p'_j\, dq'^k + ((p'_k, p'_j))_{q,p}\,\delta p'_j\, dp'_k \qquad (1.30)\end{aligned}$$

where

$$((u, v))_{q,p} = \frac{\partial q^i}{\partial u}\frac{\partial p_i}{\partial v} - \frac{\partial p_i}{\partial u}\frac{\partial q^i}{\partial v} \qquad (1.31)$$

Expression (1.31) is called the Lagrange bracket of u and v with respect to the q^i , p_i . Comparing Eqs (1.29) and (1.30), we obtain the following necessary and sufficient conditions for a canonical transformations.

$$((q'^j, q'^k))_{q,p} = 0,\quad ((p'_j, p'_k))_{q,p} = 0,\quad ((q'^j, p'_k))_{q,p} = \delta^j_k \qquad (1.32)$$

Consider now a general transformation from the variables p_i to a set of 2n variables u_μ . Introduce a matrix A whose elements are given by

$$A_{\mu\nu} = ((u_\mu, u_\nu))_{q,p} \qquad (1.33)$$

If the transformation $q, p \to u$ is non-singular (non-vanishing Jacobian) then A will non-singular, for its determinant is given by

$$|A| = \left[\frac{\partial(q,p)}{\partial(u)}\right]^2 \qquad (1.34)$$

Problem IV : Verify (1.34) Hint : Write $q^i = x_i^1$, $p_i = x_i^2$ and $((u_\mu, u_\nu))_{q,p} = \varepsilon_{\alpha\beta} \frac{\partial x_i^\alpha}{\partial u_\mu} \frac{\partial x^\beta}{\partial u_\nu}$ where $\varepsilon_{\alpha\beta}$ is the 2-dimensional alternating symbol (Levi-Civita symbol) : $(\varepsilon_{\alpha\beta}) = \begin{pmatrix} 0 & I \\ -1 & 0 \end{pmatrix}$.

The inverse of A is designated by a special notation, namely,

$$-A^{-1}_{\mu\nu} = (u_\mu, u_\nu)_{q,p}. \qquad (1.35)$$

It is readily verified that

$$(u_\mu, u_\nu)_{q,p} = \frac{\partial u_\mu}{\partial q^i}\frac{\partial u_\nu}{\partial p_i} - \frac{\partial u_\mu}{\partial p_i}\frac{\partial u_\nu}{\partial q^i} \qquad (1.36)$$

Provlem V : Verify (1.36).

That is,

$$(u_\mu, u_\sigma)_{q,p}((u_\mu, u_\sigma))_{q,p} = \delta_{\mu\nu} \qquad (1.37).$$

Expression (1.36) is called the Poisson bracket of u_μ and u_ν with respect to the q^i, p_i . Poisson brackets satisfy a set of readily verified identities. If F_1 , F_2 , F_3 are any three functions of the q^i and p_i (and perhaps also of the time), then

$$(F_1, F_2)_{q,p} = -(F_2, F_1)_{q,p} \quad , \quad (F_1, F_2 + F_3)_{q,p} = (F_1, F_2)_{q,p} + (F_1, F_3)_{q,p} \tag{1.38}$$

$$(F_1, F_2 F_3)_{q,p} = (F_1, F_2)_{q,p} F_3 + F_2 (F_1, F_3)_{q,p} \tag{1.39}$$

$$(F_1, (F_2, F_3)_{q,p})_{q,p} + (F_2, (F_3, F_1)_{q,p})_{q,p} + (F_3, (F_1, F_2)_{q,p})_{q,p} = 0 \tag{1.40}$$

Problem VI : Verify (1.39)

Problem VII : Verify (1.40). Hint : Write the expression on the left hand side as $1/2\ \varepsilon_{ijk} \left(F_i, (F_j, F_k)_{q,p}\right)_{q,p}$.

Equation (1.40) is known as the Poisson-Jacobi identity.

Using the notation of problem IV, we may write equations (1.32) in the following compact form.

$$\left((x'^{\alpha}, x'^{\beta})\right)_{q,p} = \varepsilon_{\alpha\beta}\, \delta_{ij} \tag{1.41}$$

Since $\varepsilon_{\alpha\gamma}\, \delta_{ik}\, \varepsilon_{\beta\gamma}\, \delta_{jk} = \delta_{\alpha\beta}\, \delta_{ij}$, it is evident that we also have

$$(x'^{\alpha}, x'^{\beta})_{q,p} = \varepsilon_{\alpha\beta}\, \delta_{ij} \tag{1.42}$$

or, explicitly,

$$(q'^{i}, q'^{j})_{q,p} = 0, \quad (p'^{i}, p'^{j})_{q,p} = 0, \quad (q'^{i}, p'^{j})_{q,p} = \delta_{ij} \tag{1.43}$$

Equations (1.43) may be used in place of Eqs (1.32) as a set of necessary and sufficient conditions for a canonical transformation.

One may now readily show that the variables with respect to which any Lagrange or Poisson bracket is taken are uniquely determined only up to a canonical transformation. For we have

$$((u_\mu, u_\nu))_{q,p} = \varepsilon_{\alpha\beta} \frac{\partial x_i^\alpha}{\partial u_\mu} \frac{\partial x_i^\beta}{\partial u_\nu} = \varepsilon_{\alpha\beta} \frac{\partial x_i^\alpha}{\partial x_j'^\gamma} \frac{\partial x_i^\beta}{\partial x_k'^\delta} \frac{\partial x_j'^\gamma}{\partial u_\mu} \frac{\partial x_k'^\delta}{\partial u_\nu}$$

$$= ((x_\gamma', x_\delta'))_{q,p} \frac{\partial x_j'^\gamma}{\partial u_\mu} \frac{\partial x_k'^\delta}{\partial u_\nu}$$

$$= \varepsilon_{\gamma\delta} \delta_{jk} \frac{\partial x_j'^\gamma}{\partial u_\mu} \frac{\partial x_k'^\delta}{\partial u_\nu} = ((u_\mu, u_\nu))_{q',p'} \qquad (1.44)$$

$$(u_\mu, u_\nu)_{q,p} = \varepsilon_{\alpha\beta} \frac{\partial u_\kappa}{\partial x_i^\alpha} \frac{\partial u_\rho}{\partial x_i^\beta} = \varepsilon_{\alpha\beta} \frac{\partial x_j'^\gamma}{\partial x_i^\alpha} \frac{\partial x_k'^\delta}{\partial x_i^\beta} \frac{\partial u_\mu}{\partial x_j'^\gamma} \frac{\partial u_\nu}{\partial x_k'^\delta}$$

$$= (x_\gamma', x_\delta')_{q,p} \frac{\partial u_\mu}{\partial x_j'^\gamma} \frac{\partial u_\nu}{\partial x_k'^\delta}$$

$$= \varepsilon_{\gamma\delta} \delta_{jk} \frac{\partial u_\mu}{\partial x_j'^\gamma} \frac{\partial u_\nu}{\partial x_k'^\delta} = (u_\mu, u_\nu)_{q',p'} \qquad (1.45)$$

Since Lagrange and Poisson brackets thus remain invariant under canonical transformations we shall henceforth drop the subscripts denoting the variables with respect to which they are being taken.

In addition to the above differential invariants there exist a number of integral invariants discovered by Poincaré. Consider a 2m-dimensional region R in the 2n-dimensional space of the q^i, p_i $(m \leq n)$ and consider the following integral :

$$I_m = \sum_{i_1 \ldots i_m} \int_R dq^{i_1} dp_{i_1} \ldots dq^{i_m} dp_{i_m} \qquad (1.46)$$

$i_1 \ldots i_m$ all different

If the region R is characterized by a set of 2m parameters α_κ, this integral may be rewritten in the form

$$I_m = \int_R \frac{\partial (q^{i_1} p_{i_1} \ldots q^{i_m} p_{i_m})}{\partial (\alpha_1 \ldots \alpha_{2m})} d\alpha_1 \ldots d\alpha_{2m} \qquad (1.47)$$

If we [illegible] a canonical transformation $p, p \to p', p'$ this integral becomes

$$I'_m = \int_R \frac{\partial(p'^{i_1}, p'_{i_1} \cdots p'^{i_m}, p'_{i_m})}{\partial(\alpha_1 \cdots \alpha_{2m})} d\alpha_1 \cdots d\alpha_{2m}$$

$$= \int_R \frac{1}{2^m} \varepsilon_{\alpha_1 \beta_1} \cdots \varepsilon_{\alpha_m \beta_m} \varepsilon_{\mu_1 \cdots \mu_{2m}} \frac{\partial x'^{\alpha_1}_{i_1}}{\partial \alpha_{\mu_1}} \frac{\partial x'^{\beta_1}_{i_1}}{\partial \alpha_{\mu_2}} \cdots \frac{\partial x'^{\alpha_m}_{i_m}}{\partial \alpha_{\mu_{2m-1}}} \frac{\partial x'^{\beta_m}_{i_m}}{\partial \alpha_{\mu_{2m}}} d\alpha_1 \cdots d\alpha_{2m}$$

$$= \int_R \frac{1}{2^m} \varepsilon_{\mu_1 \cdots \mu_{2m}} ((\alpha_{\mu_1}, \alpha_{\mu_2})) \cdots ((\alpha_{\mu_{2m-1}}, \alpha_{\mu_{2m}})) d\alpha_1 \cdots d\alpha_{2m}$$

$$= I_m \qquad (1.48)$$

The integral (1.46) is therefore seen to be invariant under canonical transformations.

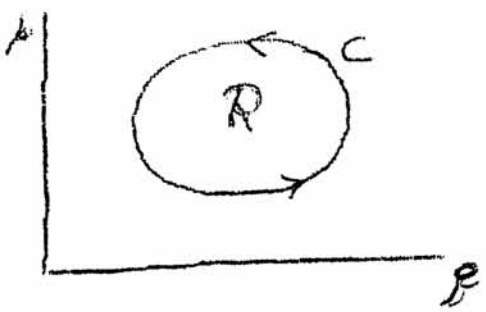
Among the above integral invariants there are two important cases :

$$1) \quad I_n = n! \int_R d\varphi^1 dp_1 \cdots d\varphi^n dp_n \qquad (1.49)$$

The 2n-dimensional spaces of the φ^i, p_i is called phase space. The invariance of I_n indicates the invariance of volume elements in phase space. (It also indicates that the Jacobian of any canonical transformation is equal to unity.)

$$2) \quad I_1 = \int_R d\varphi^i dp_i = -\oint_C p_i d\varphi^i = \oint_C p^i dp_i \qquad (1.50)$$

where C is the boundary of R and the integration is taken in the following sense.



b) Infinitesimal canonical transformations.

Suppose a canonical transformation $q, p \to q', p'$ is an infinitesimal transformation so that

$$q'^i = q^i + \Delta q^i \quad , \quad p'_i = p_i + \Delta p_i$$
$$\Delta q^i = \varepsilon \varphi^i \quad , \quad \Delta p_i = \varepsilon \psi_i \tag{1.51}$$

where the φ^i and ψ_i are certain functions of the q's and p's and ε is an infinitesimal. The canonical nature of the transformation will impose certain conditions on the functions φ^i and ψ^i. These can be obtained by writing Eq. (1.23) in the form

$$\delta V = p_i \delta q^i - p'_i \delta q'^i$$
$$= p_i \delta q^i - (p_i + \Delta p_i) \delta (q_i + \Delta q^i) = -p_i \delta \Delta q^i - \Delta p_i \delta q^i$$
$$= -\varepsilon (p_i \delta \varphi^i + \psi_i \delta q^i). \tag{1.52}$$

Now let

$$S = p_i \varphi^i + \frac{1}{\varepsilon} V. \tag{1.53}$$

Then

$$\delta S = \varphi^i \delta p_i - \psi_i \delta q^i \tag{1.54}$$

and we see that the φ^i and ψ_i must be of the form

$$\varphi^i = \frac{\partial S}{\partial p_i} \quad , \quad \psi_i = -\frac{\partial S}{\partial q^i} \tag{1.55}$$

where S is an arbitrary function of the q's and p's (and possibly also of the time).

If F is a fixed function of the canonical variables then it changes in value under the infinitesimal transformation (1.51, 55) by amount

$$\Delta F = \frac{\partial F}{\partial q^i}\Delta q^i + \frac{\partial F}{\partial p_i}\Delta p_i = \varepsilon\left(\frac{\partial F}{\partial q^i}\frac{\partial S}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial S}{\partial q^i}\right)$$
$$= \varepsilon(F,S) \tag{1.56}$$

On the other hand, if the value of F is to be regarded as being fixed, then its functional form in terms of the canonical variables must suffer a change of amount

$$\bar{\Delta} F(q,p) = F'(q,p) - F(q,p) \tag{1.57}$$

where

$$F'(q',p') = F(q,p). \tag{1.58}$$

But

$$F'(q,p) - F(q,p) = F'(q,p) - F'(q',p')$$
$$= -\Delta F'(q,p)$$
$$= -\Delta\left[F'(q,p) - F(q,p) + F(q,p)\right]$$
$$= -\Delta\bar{\Delta} F(q,p) - \Delta F(q,p) \tag{1.59}$$

and hence, to the first infinitesimal order

$$\bar{\Delta} F = -\Delta F = -\varepsilon(F,S). \tag{1.60}$$

Result (1.60) may be used to prove directly that the transformation (1.51, 55) leaves Poisson brackets invariant. We have

$$
\begin{aligned}
(F,G)_{\varphi;p'} &= \left(F(\varphi;p'),\, G(\varphi;p')\right)_{\varphi;p'} \\
&= \left(F(\varphi;p') + \Delta F(\varphi;p'),\, G(\varphi;p') + \Delta G(\varphi;p')\right)_{\varphi;p'} \\
&= \left(F(\varphi;p'), G(\varphi;p')\right)_{\varphi,p'} + \left(F(\varphi;p)\,\Delta G(\varphi;p')\right)_{\varphi;p'} \\
&\quad + \left(\Delta F(\varphi;p'), G(\varphi;p')\right)_{\varphi;p'} \\
&= (F,G)_{\varphi,p} + \Delta(F,G)_{\varphi,p} + (F,\Delta G)_{\varphi,p} + (\Delta F,G)_{\varphi,p} \\
&= (F,G)_{\varphi,p} - \varepsilon\left[\left(S,(F,G)_{\varphi,p}\right)_{\varphi,p} + \left(F,(G,S)_{\varphi,p}\right)_{\varphi,p} + \left(G,(S,F)_{\varphi,p}\right)_{\varphi,p}\right] \\
&= (F,G)_{\varphi,p}
\end{aligned}
\tag{1.61}
$$

Problem VIII : Prove that the difference between the results of applying two first order infinitesimal canonical transformations in different orders itself corresponds to an infinitesimal canonical transformation of the second order.

A given finite canonical transformation may frequently be expressible as the sum of an infinite number of infinitesimal canonical transformations. Any point in the summation may be labeled by a parameter τ, and the value of any function F of the φ's and the p's may be denoted at that point by $F(\tau)$. Then

$$\frac{dF}{d\tau} = (F,\, S) \tag{1.62}$$

where S is the generator of the transformation. S may have an explicit dependence on τ in addition to its implicit dependence on τ through its dependence on the φ's and p's.

When S is not explicitly dependent on τ, the connection between the values of F at two different points τ_1 and τ_2 may be expressed by means of the following formal expansion :

$$F_2 = F_1 + (T_2 - T_1)(F_1, S_1) + \frac{1}{2!}(T_2 - T_1)^2((F_1, S_1), S_1) + \ldots \quad (1.63)$$

The question of the actual convergence of such expansions forms one of the most difficult problems in dynamical theory.

Any canonical transformation which can be generated by infinitesimal canonical transformations is called a proper canonical transformation. The set of all proper canonical transformations forms a subgroup of the group of all canonical transformations.

Problem IX : Show that the canonical transformation

$$F' = F + T(F, S) + \frac{1}{2!}T^2((F, S), S) + \ldots$$

with $S = 1/2\ (p^i q^i + p_i p_i)$ and $T = \pi$ changes q^i into $-q^i$ and p_i into $-p_i$, and hence that inversion of coordinates is a proper canonical transformation even though for n odd it is an improper point transformation.

Problem X : Does the group of proper canonical transformations actually comprise the totality of all canonical transformations ; i.e., is the word "proper", when applied to canonical transformations redundant ?

Problem XI : Show that the canonical transformation

$$F' = F + T(F, S) + 1/2T((F, S), S) + \ldots$$

with $S = p_i \Lambda^i$ where the Λ^i are arbitrary functions of the q's (and possibly also of the time) is a proper point transformation. Each set of

---

Note - Asteriks, in general, denote the difficulty of a problem. One asterisk generally denotes a problem which can be solved readily only with the aid of outside references. Two asteriks, while by no means guaranteeing that the answer is unknown, indicates that the author is not aware of the answer and would be happy to have the answer (if any) brought to his attention.

functions Λ^i defines a one-parameter subgroup of the group of all proper point transformations. The group of all proper point transformations is itself comprised of the totatlity of these one-parameter subgroups. Show that if $\Lambda^i = \varphi^i$ and $T = \log\lambda$ then the tranformation is one which merely changes the coordinates by the scale factor λ .

Problem XII : Show that the difference between the results of applying two first order infitesimal point transformations in different orders corresponds to an infinitesimal point transformation of the second order.

Returning now the canonical equations (1.5) we see that the time derivative of any function F of the φ^i's and p_i's (and possibly also of the time) of a dynamical system is given by

$$\dot{F} = \frac{\partial F}{\partial \varphi^i}\dot{\varphi}^i + \frac{\partial F}{\partial p_i}\dot{p}_i + \frac{\partial F}{\partial t}$$
$$= (F,H) + \frac{\partial F}{\partial t} \qquad (1.64)$$

The case in which F is not explicitly dependent on t shows us, upon comparison with Eq. (1.62), that the course of development in time of a Hamiltonian system may be regarded as the gradual self-unfolding of a canonical transformation. The Hamiltonian function is the generator of the transformation.

Under the infinitesimal transformation (1.51, 55) we may write

$$p_i\delta\varphi^i - H\delta t = \left(p'_i + \epsilon\frac{\partial S}{\partial \varphi^i}\right)\delta\left(\varphi'^i - \epsilon\frac{\partial S}{\partial p_i}\right) - H\delta t$$
$$= p'_i\delta\varphi'^i + \epsilon\left(\frac{\partial S}{\partial \varphi^i}\delta\varphi^i - p_i\delta\frac{\partial S}{\partial p_i}\right) - H\delta t$$
$$= p'_i\delta\varphi'^i - \left(H + \epsilon\frac{\partial S}{\partial t}\right)\delta t + \delta V \qquad (1.65)$$

where $V = \varepsilon (S - p_i \frac{\partial S}{\partial p_i})$ (see (1.53)). Comparing with (1.22) we see that the transformed Hamiltonian function is

$$H' = H + \varepsilon \frac{\partial S}{\partial t} . \tag{1.66}$$

Suppose the form of the Hamiltonian function remains invariant under the infinitesimal transformation. Then

$$H(q', p', t) = H(q, p, t) + \varepsilon \frac{\partial}{\partial t} S(q, p, t)$$

or

$$\varepsilon \frac{\partial S}{\partial t} = \Delta H = \varepsilon (H, S) \tag{1.67}$$

Since ε is arbitrary we have

$$\dot{S} = (S, H) + \frac{\partial S}{\partial t} = 0 \tag{1.68}$$

That is, S is an integral of the dynamical system :

$$S = \text{constant} . \tag{1.69}$$

The converse of the above result is obviously also true, namely, that the canonical transformation generated by any integral of the system leaves the form of the Hamiltonian function invariant. Since the form of the Hamiltonian remains invariant, the q''s and p''s satisfy the same equations of motion as the q's and p's, and it is evident that the canonical transformation transforms trajectories into trajectories.

Let S and T be any two integrals of a dynamical system. Let the function S generate an infinitesimal canonical transformation, which transforms trajectories into trajectories. The change in the function T under such a transformation from one trajectory to another is $\Delta T = \varepsilon (T, S)$. But since, T, being an integral of the system, is constant along both trajectories, it is evident that ΔT and hence (S,T) is a constant. That is, (S,T) is also an integral of the system.

Problem XIII : Prove the above result directly.

If the Hamiltonian function is not explicitly dependent on t (this will be the case if the Lagrangian function is not explicitly dependent on t) then

$$\dot{H} = (H, H) = 0 \tag{1.70}$$

and the Hamiltonian is seen to be an integral of the system — the energy integral. Equation (1.70) expresses the law of conservation of energy, the value of the Hamiltonian being identified with the energy of the system. The canonical transformations generated by a constant Hamiltonian effect displacements of trajectories in time.

c) Action. Angle and action variables.

Let

$$\rho^i = \bar{\rho}^i(t) \tag{1.71}$$

be the solution of the equations of motion (1.1) which corresponds to the trajectory which passes through the point ρ'^i at the time t' and the point ρ^i at the time t. (There may be more than one such trajectory, but for the present we shall consider the case in which there is only one.) Then consider the following function

$$S(\rho t / \rho' t') = \int_{t'}^{t} L(\bar{\rho}, \dot{\bar{\rho}}, t)\, dt \tag{1.72}$$

S is known as the action, and is completely determined by the ρ'^i, t', ρ^i, t.

Suppose now that we make an infinitesimal vairation $\delta\bar{\rho}$ which carries us from one physical trajectory to another, so that

$$\rho^i = \bar{\rho}^i(t) + \delta\bar{\rho}^i(t) \tag{1.73}$$

is also a solution of the equations of motion. It is evident that

$$\left.\begin{aligned}\delta\rho'^i &= \delta\bar{\rho}^i(t') + \dot{\bar{\rho}}^i(t')\delta t' \\ \delta\rho^i &= \delta\bar{\rho}^i(t) + \dot{\bar{\rho}}^i(t)\delta t\end{aligned}\right\} \tag{1.74}$$

and the variation in the action becomes

$$\begin{aligned}\delta S &= L(\bar{\rho}(t), \dot{\bar{\rho}}(t), t)\delta t - L(\bar{\rho}(t'), \dot{\bar{\rho}}(t'), t')\delta t' \\ &\quad + \int_{t'}^{t}\left[\frac{\partial L(\bar{\rho},\dot{\bar{\rho}},t)}{\partial\bar{\rho}^i}\delta\bar{\rho}^i + \frac{\partial L(\bar{\rho},\dot{\bar{\rho}},t)}{\partial\dot{\bar{\rho}}^i}\delta\dot{\bar{\rho}}^i\right]dt \\ &= \bar{L}(t)\delta t - \bar{L}(t')\delta t' + \left[\frac{\partial\bar{L}}{\partial\dot{\bar{\rho}}^i}\delta\bar{\rho}^i\right]_{t'}^{t} \\ &= \left[\bar{L}(t) - \bar{p}_i(t)\dot{\bar{\rho}}^i(t)\right]\delta t - \left[\bar{L}(t') - \bar{p}_i(t')\dot{\bar{\rho}}^i(t')\right]\delta t' \\ &\quad + \bar{p}_i(t)\delta\rho^i - \bar{p}_i(t')\delta\rho'^i \\ &= -H(\rho,p,t)\delta t + H'(\rho',p',t')\delta t' + p_i\delta\rho^i - p'_i\delta\rho'^i\end{aligned} \tag{1.75}$$

where $$p_i = \bar{p}_i(t), \quad p'_i = \bar{p}_i(t') \tag{1.76}$$

Evidently

$$\frac{\partial S}{\partial\rho^i} = p_i, \quad \frac{\partial S}{\partial\rho'^i} = -p'_i \tag{1.77}$$

$$\frac{\partial S}{\partial t} = -H \qquad \frac{\partial S}{\partial t'} = H' \tag{1.78}$$

p_i and p'_i are implicit functions of the ρ'^i, t', ρ^i, t, and the action is seen to satisfy the equations

$$\frac{\partial S}{\partial t} + H\left(\rho, \frac{\partial S}{\partial\rho}, t\right) = 0 \tag{1.79}$$

Action is seen to satisfy the equations

$$-\frac{\partial S}{\partial t'} + H\left(q', -\frac{\partial S}{\partial q'}, t\right) = 0 \qquad (1.80)$$

These are known as the Hamilton-Jacobi equations.

Suppose that, instead of the q^i and q'^i being given, we know only the initial values, q'^i, p'_i, of the dynamical variables at the time t'. Then a knowledge of the function $S(q, t / q', t')$ would enable us to determine the values q^i, p_i at any later (or earlier) time t. To see how this comes about, first recall that the development of a dynamical system in time corresponds to the self-unfolding of a canonical transformation. Comparison of Eq. (1.18) with (1.77) shows that the latter is an explicit expression of this transformation. (S is to be identified with the Ω_1, of (1.18)). Equations (1.77) may evidently be solved for the q^i and p_i in terms of t, t' and the initial values q'^i, p'_i:

$$q^i = q^i(q', p', t', t) \quad , \quad p_i = p_i(q', p', t', t) \qquad (1.81)$$

In this manner, if a general solution of the Hamilton-Jacobi equation (1.79) is found, the equations of motion are immediatly solved.

The function S is a particular form of general solution of the Hamilton-Jacobi equation. There exist many other useful forms for the general solution. For example, solve equations (1.77) for the q'^i in terms of the q^i, p'_i, t and t',

$$q'^i = q'^i(q, p', t, t'), \qquad (1.82)$$

and then introduce the function

$$T(q, t / p', t') = S(q, t / q'(\ ), t') + q'^i(\ ) p'_i . \qquad (1.83)$$

T may be identified with the Ω_2 of (1.19), and the canonical transformation (1.77) may be reexpressed in the form

$$\frac{\partial T}{\partial q^i} = p_i \quad , \quad \frac{\partial T}{\partial p'_i} = q'^i . \qquad (1.84)$$

We also have

$$\frac{\partial T}{\partial t} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial \rho'^i}\frac{\partial \rho'^i}{\partial t} + p'_i\frac{\partial \rho'^i}{\partial t} = \frac{\partial S}{\partial t} = -H, \tag{1.85}$$

and hence

$$\frac{\partial T}{\partial t} + H\left(\rho, \frac{\partial T}{\partial \rho}, t\right) = 0 \tag{1.86}$$

Problem XIV : Prove that T also satisfies the partial differential equation

$$-\frac{\partial T}{\partial t'} + H\left(\frac{\partial T}{\partial \rho'}, \rho', t'\right) = 0$$

An infinity of other forms for the general solution may be obtained by carrying out a canonical transformation $\rho', p' \rightarrow \rho'', p''$ on the ρ''s and p''s. Let the canonical transformation be generated by a function $\Omega_4(\rho', p'')$, for example, so that, according to (1.19),

$$\rho'^i = -\frac{\partial \Omega_4}{\partial p'_i}, \quad p''^i = \frac{\partial \Omega_4}{\partial \rho''^i} \tag{1.87}$$

The first of Eqs. (1.87) may be combined with the second of Eqs. (1.84) to solve for the p'_i in terms of the ρ^i, p''_i, t. (t' also occurs, but we shall now let its presence be merely implicitly understood).

$$p'_i = p'_i(\rho, p'', t). \tag{1.88}$$

Then the canonical transformation which passes from the ρ's and p's to the ρ'''s and p'''s is generated by the function

$$U(\rho, p'', t) = T(\rho, t/p'(), t') + \Omega_4(p'(), p'') \tag{1.89}$$

We have

$$\frac{\partial U}{\partial p^i} = p_i \quad , \quad \frac{\partial U}{\partial p''_i} = p''^i . \tag{1.90}$$

Now, under the canonical transformation $p, p \rightarrow p'', p''$, the Hamiltonian function transforms according to

$$H'' = H + \frac{\partial U}{\partial t} \tag{1.91}$$

But

$$\frac{\partial U}{\partial t} = \frac{\partial T}{\partial t} + \frac{\partial T}{\partial p'_i}\frac{\partial p'_i}{\partial t} + \frac{\partial \Omega_4}{\partial p'_i}\frac{\partial p'_i}{\partial t}$$

$$= \frac{\partial T}{\partial t} + (p'^i - p'^i)\frac{\partial p'_i}{\partial t} = \frac{\partial T}{\partial t} = -H \tag{1.92}$$

Hence

$$H'' = 0 \tag{1.93}$$

and the p'''s and p'''s satisfy the equations

$$p''^i = 0, \quad p''_i = 0 . \tag{1.94}$$

That is to say, the p'''s and p'''s are constants as far as variation with t is concerned. This is actually already obvious, since the p'''s and p'''s are simply functions of the p''s and p''s, which are in turn independent of t. Equations (1.90) may be solved for the p's and p's in terms of t and the constants p''^i, p''_i:

$$p^i = p^i(p'', p'', t) \quad , \quad p_i = p_i(p'', p'', t) \tag{1.95}$$

The relation between the constants p''^i, p''_i and the initial values p'^i, p'_i may be obtained by setting $t = t'$ in (1.95).

Now, in virtue of Eq. (1.92) it is evident that

$$\frac{\partial U}{\partial t} + H\left(q, \frac{\partial U}{\partial q}, t\right) = 0. \tag{1.96}$$

Expression (1.89) is therefore a solution of the Hamilton-Jacobi equation, and this is the form of general solution which we shall employ in the theoretical discussion which follows.

We may first make some remarks concerning the nature of the general solution of the Hamilton-Jacobi equation which should be expected. The Hamilton-Jacobi equation is a partial differential equation of the first order in the n + 1 independent variables q^i, t. A general solution must therefore involve n + 1 independent constants of integration. Since only derivatives of U occur in (1.96), one of these constants of integration must be a purely additive constant. Now, it will be noted that U as given by (1.89) differs in value from S only by a constant, namely $q'^i p'_i + \Omega_4$. We will write henceforth $U = S$ + constant, where the constant is assumed to be adjustable. The remaining n constants of integration are the p''^i in (1.89).

Now suppose that the Hamiltonian function is not explicitly dependent on the time. Then the solution of equation (1.96) may be expressed in the form

$$U(q, p'', t) = W(q, p'') - E(p'')t \tag{1.97}$$

where the function W, which is independent of the time, satisfies the partial differential equation

$$H\left(q, \frac{\partial W}{\partial q}\right) = E(p''). \tag{1.98}$$

This is known as the second form of the Hamilton-Jacobi equation.

Introduce the further very simple canonical transformation $q'', p'' \rightarrow q''', p'''$, where

$$q'''^i = \Omega^i t + q''^i \quad , \quad \Omega^i = \frac{\partial E}{\partial p''_i} \tag{1.99}$$

$$p'''_i = p''_i \tag{1.100}$$

It is readily seen that the canonical transformation $q, p \to q''', p'''$ is generated by the function W :

$$p_i = \frac{\partial U}{\partial q^i} = \frac{\partial W}{\partial q^i} \tag{1.101}$$

$$q'''^i = \frac{\partial U}{\partial p''_i} + \frac{\partial E}{\partial p''_i} t = \frac{\partial W}{\partial p'''_i} \tag{1.102}$$

The value of W is readily obtained

$$W = U + Et = S + Ht + \text{const.}$$

$$= \int_{t'}^{t} (L + H)\, dt + Ht' + \text{const.} = \int_{q'}^{q} p_i\, dq^i + Ht' + \text{const.} \tag{1.103}$$

The constant may be adjusted so that

$$W = \int_{q'}^{q} p_i\, dq^i \tag{1.104}$$

W is sometimes also called the action. Introducing a momentum $\lambda = -H$ conjugate to the time t, one will note that we may write (since $p_i \dot{q}^i - H = L$)

$$S = \int_{\text{trajectory}} (p_i\, dq^i + \lambda\, dt) \tag{1.105}$$

This suggests a simple distinction between S and W. W may be called the space-action while S is called the space-time action.

If, in Eq. (1.98), p''_1 (or p'''_1) is chosen equal to E, then $\Omega^1 = 1$ and $\Omega^i = 0$ for $i = 2 \ldots n$. $q'''^2 \ldots q'''^n$ are then constant in time and the last n-1 of Eqs. (1.102) may be solved to express n-1 of the coordinates q^i in terms of a remaining coordinate. In this way we get the trajectories (or orbits), in q^i-space directly without the necessity of expressing the q^i separately as functions of the time.

Problem : For the harmonic oscillator

$$L = \frac{m}{2}\dot{q}^2 - \frac{mc\omega^2}{2}q^2$$

show that $S(q,t/q',t') = \frac{m\omega}{2}\frac{(q^2+q'^2)\cos\omega(t-t') - 2qq'}{\sin\omega(t-t')}$

and $T(q,t/p',t') = -\frac{1}{\omega}\frac{\sin\omega(t-t')}{\cos\omega(t-t')}\left[\frac{1}{2m}p'^2 + \frac{m\omega^2}{2}q^2\right] + \frac{qp'}{\cos\omega(t-t')}$

Show that S and T satisfies the Hamilton-Jacobi equations with

$$H = \frac{1}{2m}p^2 + \frac{m\omega^2}{2}q^2$$

Show that $W(q,E) = -\frac{m\omega}{2}\left[q\sqrt{\frac{2E}{m\omega^2} - q^2} - \frac{2E}{m\omega^2}\cos^{-1}\left(\sqrt{\frac{m\omega^2}{2E}}\,q\right)\right]$

+ constant

(Some signs are arbitrary.)

Problem : For the freely falling body

$$L = \frac{m}{2}\dot{q}^2 - mgq \quad , \quad H = \frac{1}{2m}p^2 + mgq$$

show that

$$S = \frac{m}{2}\frac{(q-q')^2}{t-t'} - \frac{m}{2}g(q+q')(t-t') - \frac{mg^2}{24}(t-t')^3$$

$$T = -\left[\frac{1}{2m}p'^2 + mgq\right](t-t') + qp' + \frac{g}{2}p'(t-t')^2 - \frac{mg^2}{6}(t-t')^3$$

$$W = \pm\frac{1}{3m^2g}(2mE - 2m^2gq)^{3/2} + \text{constant}$$

Problem : Show that for a charged body moving in a plane in a constant homogeneous magnetic field H perpendicular to the plane

$$L = \frac{m}{2}\dot{\rho}^i\dot{\rho}^i + \frac{e}{c}A_i\dot{\rho}^i$$

$$= m\left(\frac{1}{2}\dot{\rho}_i\dot{\rho}_i + \frac{\omega}{2}\varepsilon_{ij}\rho^i\dot{\rho}^j\right) \qquad i=1,2$$

$$\omega = \frac{eH}{mc}, \quad A_i = \frac{1}{2}H\varepsilon_{ij}\rho^j$$

$$S = \frac{m\omega}{4}\frac{(\rho^*-\rho'^*)(\rho-\rho')\sin\omega(t-t')}{1-\cos\omega(t-t')} + i\frac{m\omega}{4}(\rho^*\rho'-\rho'^*\rho)$$

$$T = -\frac{A}{m\omega}\frac{1-\cos\omega(t-t')}{\sin\omega(t-t')}\left(p'+\frac{im\omega}{4}\rho^*\right)\left(p'^*-\frac{im\omega}{4}\rho\right)+p\rho+p^*p'^*$$

where $\rho = \rho' + i\rho^2$, $p = \frac{1}{2}(p_1 - ip_2)$

$(\rho, p) = 1$ $\quad (\rho^*, p^*) = 1$ $\quad (\rho, p^*) = 0$ etc.

It will be observed that the space-action W for the two-dimensional harmonic oscillator is expressible in the form

$$W = \sum_{i=1}^{2} W_i \qquad (1.141)$$

where W_i is a function only of ρ^i, together with the integration constants. This is obviously a direct consequence of the fact that the Hamiltonian function has the form

$$H = \sum_{i=1}^{2} H_i \qquad (1.142)$$

where H_i is a function of ρ^i and p_i only. In this case the Hamilton-Jacobi equation (1.98) breaks up into the equations

$$H_i\left(q^i, \frac{\partial W}{\partial q^i}\right) = p_i'' \quad , NS, \quad (i=1,2) \tag{1.143}$$

It may happen that the space-action is expressible in the form

$$W(q, p'') = \sum_{i=1}^{n} W_i(q^i, p'') \tag{1.144}$$

even when the Hamiltonian does not consist of a sum of terms each of which depends on only one pair of variables q^i, p_i. Whenever W is expressible in the form (1.144), the Hamilton-Jacobi equation (1.98) is said to be separable.

The exact solubility in closed form of the equations of motion of many simple systems is often directly related to the fact that the Hamilton-Jacobi equation for the system in question is separable. In the discussion immediatly following we shall consider exclusively systems for which the Hamilton-Jacobi equation is separable. Equations (1.101, 102) will then take the forms

$$p_i = \frac{\partial W_i}{\partial q^i} \qquad NS. \tag{1.145}$$

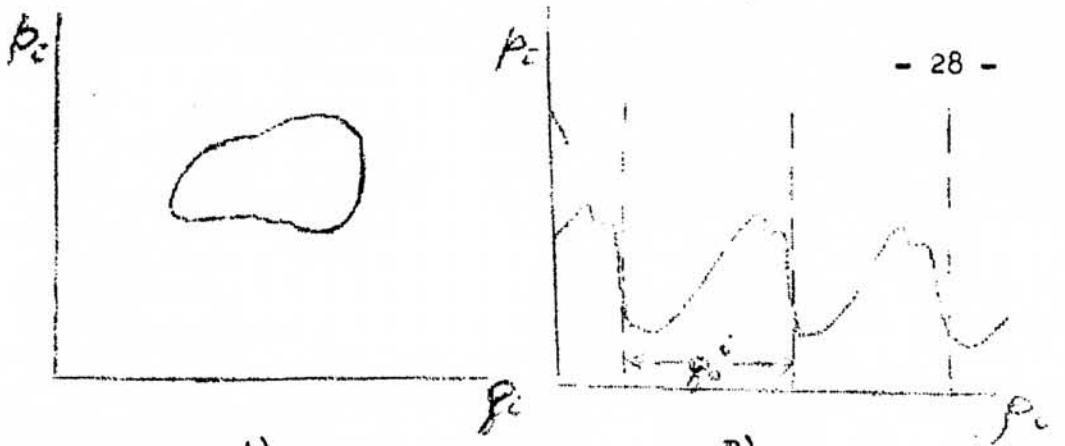
$$q'''^i = \sum_{\delta=1}^{n} \frac{\partial W_\delta}{\partial p_i'''} = \sum_{\delta=1}^{n} \frac{\partial W_\delta}{\partial p_i''} \tag{1.146}$$

and, by solving these equations, we may write

$$q^i = q^i(q''', p''') = q^i(\Omega t + q'', p'') \tag{1.147}$$

$$p_i = p_i(q^i, p'') \tag{1.148}$$

Now suppose that the system is of such a kind that canonical variables q^i, p_i can be introduced for which the projections of the phase-space orbit of the system on all the n $q - p$ planes, as given by equations (1.148), have one or other of the two following forms, for a certain range of values of the integration constants p'' :



A) B)

The motion of the system is then said to be multiply periodic in this range. If the ith projection has the form A), the $\rho^i - p_i$ motion is said to be one of libration. If it has the form B) then the $\rho^i - p_i$ motion is said to be one of rotation. The significance of the "wavelength" ρ_0^i in case B) must always be understood in the following sense. If the coordinate ρ^i is increased by amount ρ_0^i while all the others remain fixed, then the configuration of the system remains unchanged. By carrying out the point transformation

$$Q^i = \sin\left(2\pi \frac{\rho^i}{\rho_0^i}\right) \quad NS \tag{1.149}$$

$$P^i = \frac{\rho_0^i}{2\pi} \frac{p_i}{\cos(2\pi \rho^i/\rho_0^i)} \quad NS \tag{1.150}$$

the case of rotation can be transformed into a case of libration. We shall therefore henceforth consider the case of libration only.

Now introduce the quantities

$$J_i = \frac{1}{2\pi} \oint \frac{\partial W_i(\rho^i, \rho'')}{\partial \rho^i} d\rho^i = \frac{1}{2\pi} \oint p_i d\rho^i \quad NS \tag{1.151}$$

where the integration is to be carried over one complete $\rho^i - p_i$ cycle. It is evident that

$$J_i = \frac{\Delta W_i}{2\pi} \tag{1.152}$$

where ΔW_i denotes the increase in the i th component of the space action

W during one cycle. It is for this reason that the J_i are commonly known as action variables. It should be noted that our definition here differs from the usual one, namely $J_i = \Delta W_i$, by a factor 2π. It will be observed from (1.151) that the J_i are functions of the p''_i alone, and are hence constants. Solving for the p'''s in terms of the J's, one may reexpress the space action W as a function of the q's and the J's. W then generates a canonical transformation from the q's and p's to a set of variables w^i, J_i :

$$p_i = \frac{\partial W}{\partial q^i} \quad , \quad w^i = \frac{\partial W}{\partial J_i} \quad . \tag{1.153}$$

The canonical transformation which leads from the q''''s and p''''s (or from the q'''s and p'''s) to the w's and J's is simply a point transformation in momentum space,

$$J_i = J_i(p''') = J_i(p'') \tag{1.154}$$

$$w^i = \frac{\partial q'''_i}{\partial J_i} q'''^i = \omega^i t + \varphi^i, \tag{1.155}$$

where

$$\omega^i = \frac{\partial p''_i}{\partial J_i} \Omega_j \quad , \quad \varphi^i = \frac{\partial p''_i}{\partial J_i} q''^j . \tag{1.156}$$

The change in the variable w^i during one cycle of the coordinate q^j and momentum p_j is given by

$$\Delta_j w^i = \oint \frac{\partial w^i}{\partial q^j} dq^j = \oint \frac{\partial^2 W}{\partial q^j \partial J_i} dq^j \qquad NS$$

$$= \frac{\partial}{\partial J_i} \oint \frac{\partial W_j}{\partial q^j} dq^j = 2\pi \frac{\partial J_j}{\partial J_i} = 2\pi \delta^i_j \qquad NS \tag{1.157}$$

That is, w^i remains unchanged over cycles j with $j \neq i$ but changes

by an amount 2π over the cycle i. For this reason the variables ω^i, which are canonically conjugated to the action variables J_i, are called angle variables (winkelvariabeln). Conversely, if we allow ω^i to increase by an amount 2π, p^i will go through a complete cycle and then return to its initial value. The other p's may also depend on ω^i but will return to their initial values without going through a complete cycle. It is evident therefore that the general solutions of equations (1.153) have the forms

$$p^i = \sum_{\tau_j=-\infty}^{\infty} p^i_{\tau_1 \dots \tau_n} e^{i\tau_j \omega^j}, \quad p_i = \sum_{\tau_j=-\infty}^{\infty} p_{i\,\tau_1 \dots \tau_n} e^{i\tau_j \omega^j} \tag{1.158}$$

If we choose either the zero point of time or the zero point of the angle variables in such a way that $\varphi^i = 0$ in (1.155), we may write

$$p^i = \sum_{\tau_j=-\infty}^{\infty} p^i_{\tau_1 \dots \tau_n} e^{i\tau_j \omega^j t}, \quad p_i = \sum_{\tau_j=-\infty}^{\infty} p_{i\,\tau_1 \dots \tau_n} e^{i\tau_j \omega^j t} \tag{1.159}$$

From the reality condition on the p's and p's we must have

$$p^{i\,*}_{\tau_1 \dots \tau_n} = p^i_{-\tau_1 \dots -\tau_n}, \quad p^*_{i\,\tau_1 \dots \tau_n} = p_{i\,-\tau_1 \dots -\tau_n} \tag{1.160}$$

as a condition on the amplitudes of the multiple Fourier series.

From equation (1.159) the ω^i are seen to be the angular frequencies of the multiply periodic motion. Since the total energy of the system is a function of the constants p_i" it is also a function of the action variables J_i. From equations (1.99) and (1.156) it is evident that the angular frequencies are given by

$$\omega^i = \frac{\partial E}{\partial J_i}, \tag{1.161}$$

The ω^i wil in general depend on the J's.

The canonical transformation equations (1.153) correspond to the form II) on page 5, W here taking the place of the generating function Ω_2. Of interest is also another generating function W for the same

canonical transformation, namely that which corresponds to the form I) on page 5 :

$$w = W - \omega^i J_i, \tag{1.162}$$

$$p_i = \frac{\partial w}{\partial q^i}, \quad J_i = -\frac{\partial w}{\partial \omega^i} \tag{1.163}$$

As ω^i increases byan amount 2π, q^i will go through a complete cycle while the remaining q's do not. Hence W changes by amount ΔW_i, while w changes by amount

$$\Delta w = \Delta W_i - 2\pi J_i = 0, \tag{1.164}$$

That is, w is evidently periodic in the ω's with period 2π.

We shall now generalize the foregoing analysis of mechanical systems in terms of angle and action variables, in such a way that no reference need be made to the possible separability of the Hamilton-Jacobi equation. Any system, whose Hamiltonian function is not explicitly dependent upon the time, for whicha canonical transformation $q,p \to \omega, J$ of the form (1.163) can be found, such that :
1) the Hamiltonian function transforms into a function E depending only on the J's, 2) the q's and p's are periodic in the ω's with period 2π, and 3) the generating function w is also periodic in the ω's with period 2π, both through its explicit dependence on the ω's as well as implicitly through its dependence on the q's, will be called a multiply-periodic system. The ω's and J's will be called the angle and action variables of the system. We must clearly investigate the question of the uniqueness of these angle and action variables.

Consider the following linear point transformation of the angle and action variables :

$$\omega'^i = \tau^i_j \omega^j, \quad J'_i = \tau^{-1}{}^j_i J_j, \tag{1.165}$$

with

$$\tau^{-1}{}^{i}_{k}\,\tau^{k}_{j} = \delta^{i}_{j}, \tag{1.66}$$

where the τ^{i}_{j} are constant integral coefficients. If the determinant of the matrix formed by the τ^{i}_{j} has the value

$$|\tau^{i}_{j}| = \pm 1 \tag{1.167}$$

then the elements $\tau^{-1}{}^{i}_{j}$ of the inverse matrix are also integers. This means that not only does an increase in any one of the ω^{i} by amount 2π bring about increments in the ω'^{i} which are integral multiples of 2π, but conversely also an increment of 2π in one of the ω'^{i} corresponds to increments in the ω^{i} which are integral multiples of 2π. Hence any function which is periodic in the ω^{i} with period 2π is also periodic in the ω'^{i} with period 2π, and vice versa. The variables ω'^{i}, J'^{i} clearly satisfy all the requirements of angle and action variables. The J'^{i}s are constants and are integrally related to the J^{i}s. The ω'^{i}s like the ω^{i}s, are linearly dependent on the time.

It is therefore evident that the angle and action variables are unique at most up to a linear tranformation with determinant equal to ∓ 1. We shall presently see, however, that they may not even be determined to this degree of uniqueness.

It frequently happens that there exists a linear relationship between the angular frequencies ω^{i} (as defined by (1.161) of the form

$$\tau_{i}\omega^{i} = 0 \tag{1.168}$$

where the τ_{i} are integers independent of the J's. We shall suppose that equation (1.168) has been reduced to its lowest common denominator, that is, that the τ_{i} have no common integral factor other than ∓ 1. We shall further suppose that the ω^{i} have been arranged in such an order that τ_{1} and τ_{2} are different from zero and contain no common integral factor other than ∓ 1.* It is then always possible to choose integers τ'_{i}

* If only one of the τ_i, say τ_1, is different from zero, then we have immediately $\omega' \equiv 0$

---

and τ_2' such that

$$\begin{vmatrix} \tau_1 & \tau_2 \\ \tau_1' & \tau_2' \end{vmatrix} = \pm 1. \tag{1.169}$$

If we now make the point transformation

$$\left.\begin{aligned} \omega'^1 &= \tau_i \omega^i \\ \omega'^2 &= \tau_1' \omega^1 + \tau_2' \omega^2 \\ \omega'^i &= \omega^i \quad \text{for } i = 3, 4 \ldots n \end{aligned}\right\} \tag{1.170}$$

then the angular frequency ω'^1 corresponding to the new angle variable ω'^1 vanishes, and ω'^1 itself is seen to be constant in time.

Again it may happen that there exists an integral linear relationship between the new frequencies ω'^i of the form

$$\tau_i'' \omega'^i = 0 \tag{1.171}$$

where in the summation, the index i now runs over the values 2, 3,...n. Introducing the point transformation

$$\left.\begin{aligned} \omega''^1 &= \omega'^1 \\ \omega''^2 &= \tau_2'' \omega'^2 + \tau_3'' \omega'^3 + \ldots + \tau_n'' \omega'^n \\ &\text{etc.} \end{aligned}\right\} \tag{1.172}$$

we obtain another vanishing angular frequency. Continuing in this way, we may separate the angle and action variables into two groups, ω^A, J_A, and ω^μ, J_μ, such that the angular frequencies corresponding to the variables ω^A, J_A vanish, while those corresponding to the variables ω^μ, J_μ are all incommensurable with one another.

If the number of the variables ω^A is m, then the system is said to be m-fold degenerate. From the equations

$$\omega^A = \frac{\partial E}{\partial J_A} = 0 \tag{1.173}$$

it is evident that the energy is independent of the action variables J_A. It may occasionally happen that the derivative $\partial E / \partial J_i$ vanishes for

some ℓ for certain values of the J's. The system will then be said to be accidentally degenerate for those values of the J's. True degeneracy is independent of the values of the J's and implies that the energy does not depend at all on certain of the J's.

From now on, we shall maintain the above separation of the angle and action variables into a degenerate and a non degenerate group. The capital Latin indices A, B, etc.. will be assumed to range over the values 1, 2.... m, while the Greek indices range over the values m+1...n. The transformed Hamiltonian function E (i.e. the energy) will always depend on exactly n-m non-degenerate action variables J_ν . This means that the point transformations (1.165) must be restricted to those having the form

$$\left.\begin{aligned} w'^A &= \tau^A{}_B\, w^B, \quad w'^\mu = \tau^\mu{}_\nu\, w^\nu \\ J'_A &= \tau^{-1}{}^B{}_A\, J_B, \quad J'_\mu = \tau^{-1}{}^\nu{}_\mu\, J_\nu \end{aligned}\right\} \qquad (1.174)$$

where we have separately

$$|\tau^A{}_B| = \pm 1, \qquad |\tau^\mu{}_\nu| = \pm 1. \qquad (1.175)$$

We shall now show that the transformations (1.174) are too restrictive, and that there exists a wider class of canonical transformations which lead to variables satisfying all the requirements of angle and action variables. This wider class consists, namely, of all those transformations generated by functions Ω_2 having the form

$$\Omega_2(w, J') = J'_A\, \tau^A{}_B\, w^B + J'_\mu\, \tau^\mu{}_\nu\, w^\nu + F(w^\beta, J') \qquad (1.176)$$

where the function F is periodic in the w^β with period 2π. The τ_{AB}, $\tau^\mu{}_\nu$, of course, are integers satisfying (1.175) The transformation generated by Ω_2 is given by

$$\left.\begin{aligned} w'^A &= \tau^A{}_B\, w^B + \frac{\partial E}{\partial J'_A}(w^\beta, J') \\ w'^\mu &= \tau^\mu{}_\nu\, w^\nu + \frac{\partial E}{\partial J'_\mu}(w^\beta, J') \end{aligned}\right\} \qquad (1.177)$$

$$\left.\begin{aligned} J_A &= J'_B \tau^B_A + \frac{\partial F}{\partial w^A}(w^B, J') \\ J_\mu &= J'_\nu T^\nu_\mu \end{aligned}\right\} \qquad (1.178)$$

It is easily seen that the transformed function W' which generates the canonical transformation $q, p \to w', J'$ is given by

$$W'(q, J') = W(q, w(q, J')) + \Omega_2(w(q, J'), J') \qquad (1.179)$$

where the equations $J_i = -\partial W / \partial w^i = \partial \Omega_2 / \partial w^i$ have been solved to express the w^i's as functions of the q^i's and J'^i's. W', like W, satisfies the Hamilton-Jacobi equation.

Now, equations (1.177, 178) show that the w^i are changed by integral multiples of 2π when any one of the w^i is changed by 2π, and vice versa. Hence any function which is periodic in the w^i is also periodic in the w'^i. The J'^i's are evidently unaffected by increments of 2π in the w^i, although the J'_A may be altered by changes in the w^A of magnitude less than 2π. Since the w^A are independent of the time however, the J'_A are still constants. The J'_μ are, of course, constants. The transformed Hamiltonian function remains explicitly dependent solely on these non-degenerate action variables. It only remains to show that the function W' which generates the canonical transformation $q, p \to w', J'$ in the form (1.163) is periodic in the w'^i's.

Regarded as a function of the w'^i's and J'^i's, Ω_2 has the form

$$\Omega_2 = J'_A w'^A + J'_\mu w'^\mu - J'_A \frac{\partial F}{\partial J'_A} - J'_\mu \frac{\partial F}{\partial J'_\mu} + F \qquad (1.180)$$

Hence

$$\begin{aligned} W' &= W - w'^A J'_A - w'^\mu J'_\mu \\ &= W + \Omega_2 - w'^A J'_A + w'^\mu J'_\mu \\ &= W - J'_A \frac{\partial F}{\partial J'_A} - J'_\mu \frac{\partial F}{\partial J'_\mu} + F \end{aligned} \qquad (1.181)$$

Since W and F are periodic in the w^k's, and hence also in the ω^k's with the period 2π, it is evident that $2W^1$ is also.

It is fairly easy to see that the canonical transformations (1.177, 178) are the most general transformations which maintain the separation of the angle and action variables into the degenerate and non-degenerate groups. The function F (and hence its derivatives) must be periodic in the w^s, for otherwise W^1 would not be. F cannot depend on the ω^k, for otherwise the J_k would vary with time. It will be noted that the non-degenerate action variables can only transform linearly and integrally.

Problem : For the harmonic oscillator show that

$$W = -\frac{m\omega}{2} q^2 \cot \omega$$
$$= -\frac{1}{2} J \sin 2\omega$$

Problem XIX : Show that by carrying out the canonical transformation generated by the function

$$S_1 = m\omega\left(-\frac{1}{2} q^1 q^2 + q^1 Q^2 + q^2 Q^1 - Q^1 Q^2\right)$$

one can transform the Hamiltonian function of the system of Problem XVIII into that of a one-dimensional harmonic oscillator, in the form

$$H = \frac{1}{2m} P_2^2 + \frac{m\omega^2}{2} Q^2{}^2$$

Using the one-dimensional form of Eq. (1.188) show that the space action for this case takes the form

$$W = \pm \frac{m\omega}{4}\left[\frac{16J_1J_2}{m^2\omega^2} - \left(\frac{2J_1}{m\omega} - \rho'^2 + \frac{2J_2}{m\omega} - \rho^2\right)^2\right]^{1/2}$$

$$+ J_1 \cos^{-1}\frac{\sqrt{m\omega}}{2J_1}\frac{}{2(\rho'^2+\rho^2)}\left\{\rho'\left(\rho'^2+\rho^2+\frac{2J_1}{m\omega}-\frac{2J_2}{m\omega}\right)\right.$$

$$\left. \pm \rho^2\left[\frac{16J_1J_2}{m^2\omega^2} - \left(\frac{2J_1}{m\omega} - \rho'^2 + \frac{2J_2}{m\omega} - \rho^2\right)^2\right]^{1/2}\right\}$$

$$+ J_2 \cos^{-1}\frac{\sqrt{m\omega}}{2J_2}\frac{}{2(\rho'^2+\rho^2)}\left\{\rho^2\left(\rho'^2+\rho^2+\frac{2J_2}{m\omega}-\frac{2J_1}{m\omega}\right)\right.$$

$$\left. \pm \rho'\left[\frac{16J_1J_2}{m^2\omega^2} - \left(\frac{2J_1}{m\omega} - \rho'^2 + \frac{2J_2}{m\omega} - \rho^2\right)^2\right]^{1/2}\right\}$$

and that

$$E = J_2\,\omega$$

Show that the variables ω_i, J_i satisfy all the requirements of angle and action variables and hence that the system is periodic with a 1-fold degeneracy.

Problem XX : Consider the case of a one-dimensional particle in a box of length ℓ : $L = 1/2\, m\dot{x}^2 - \alpha\left(\frac{2x}{\ell}\right)^\infty$, $H = \frac{1}{2m}p^2 + \alpha\left(\frac{2x}{\ell}\right)^\infty$.

Prove that $E = \frac{\pi^2}{2m\ell^2} J^2$ and show directly that the angular frequency of oscillation is given by $\partial E/\partial J$. Show that in terms of the angle and action variables, x and p are given by

$$p = \sum_{n=-\infty}^{\infty} \frac{\ell}{n^2\pi^2}\left[1-(-1)^n\right] e^{inw},$$

$$p = \sum_{n=-\infty}^{\infty} \frac{iZ}{n\ell}\left[1-(-1)^n\right] e^{inw}.$$

Use simple Fourier analysis (See Eqs. (2.7) and (2.8) of the Appendix).

The preceding example, together with those in the two problems above, are examples of libration. As an example of rotation we may take the very simplest, though typical, case, namely the plane rotator with the Lagrangian function

$$L = 1/2\, I\dot{\varphi}^2 \tag{1.195}$$

I is the moment of inertia and φ is the angular configuration coordinate. $\varphi + 2\pi$ represents the same configuration as does φ. In order to reduce this to the case of libration we make the point transformation

$$p = \sin\varphi \quad , \quad \rho = \frac{\Pi}{\cos\varphi} \tag{1.196}$$

where Π is the momentum conjugate to φ :

$$\Pi = \frac{\partial L}{\partial \dot{\varphi}} = I\dot{\varphi}. \tag{1.197}$$

In terms of the new variables the Lagrangian function takes the form

$$L = \frac{1}{2} I \frac{\dot{p}^2}{1-p^2} \tag{1.198}$$

The Hamiltonian function has the form

$$H = \frac{1}{2I}\Pi^2 = \frac{1}{2I}(1-p^2)\rho^2. \tag{1.199}$$

The equations of motion are

$$\ddot{\varphi} = 0 \quad \text{or} \quad (1-p^2)\ddot{p} + p\dot{p}^2 = 0 \tag{1.200}$$

with solutions

$$\varphi = \omega t + \varphi_0 \quad \Rightarrow \quad \rho = \sin(\omega t + \varphi_0). \tag{1.201}$$

The second form of the Hamilton-Jacobi equation is

$$\frac{1}{2I}(1-\rho^2)\left(\frac{\partial W}{\partial \rho}\right)^2 = E \tag{1.202}$$

with the immediate solution

$$W = \sqrt{2\,I\,E}\,\sin^{-1}\rho + \text{const.} \tag{1.203}$$

Now, from the solution (1.201) we obtain

$$\pi = I\omega \quad \text{and} \quad H = \frac{1}{2} I \omega^2 = E. \tag{1.204}$$

The action variable is

$$J = \frac{1}{2\pi}\oint p\,dq = \frac{1}{2\pi}\oint \pi\,d\varphi = \pi = I\omega. \tag{1.205}$$

Hence

$$E = \frac{J^2}{2I}, \tag{1.206}$$

which chcks with

$$\omega = \frac{\partial E}{\partial J} = \frac{J}{I}. \tag{1.207}$$

We may now write

$$W = J\,\sin^{-1}\rho + \text{const.} \tag{1.208}$$

which gives

$$\omega = \frac{\partial W}{\partial J} = \sin^{-1}\rho = \varphi \tag{1.209}$$

$$p = \frac{\partial W}{\partial \varphi} = \frac{J}{\sqrt{1-\rho^2}} = \frac{J}{\cos\omega} \tag{1.210}$$

In this case the function $\mathcal{W}$ of (1.162) becomes merely a constant.

$$\mathcal{W} = W - \omega J = \text{const.} \tag{1.211}$$

$\mathcal{W}$ is therefore manifestly periodic in ω but useless as for as defining the canonical transformation $\varphi, p \to \omega, J$ is concerned.

A most important example of the use of angle and action variables is that of motion in a central field of force. The Lagrangian function is, in spherical coordinates,

$$L = 1/2\, m \left(\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2 \sin^2\theta \dot{\varphi}^2\right) - V(r) \tag{1.212}$$

The conjugate momenta are

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \tag{1.213}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = m r^2 \dot{\theta} \tag{1.214}$$

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = m r^2 \sin^2\theta\, \dot{\varphi} \tag{1.215}$$

and the Hamiltonian function is

$$E = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right) + V(r) \tag{1.216}$$

The second form of the Hamilton-Jacobi equation therefore becomes

$$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{1}{r^2 \sin^2\theta}\left(\frac{\partial W}{\partial \varphi}\right)^2\right] + V(r) = E \tag{1.217}$$

This equation is separable, for we may write

$$W(r,\theta,\varphi) = W_r(r) + W_\theta(\theta) + W_\varphi(\varphi), \tag{1.218}$$

leading to

$$\frac{dW_\varphi}{d\varphi} = a_\varphi \tag{1.219}$$

$$\left(\frac{dW_\theta}{d\theta}\right)^2 + \frac{a_\varphi^2}{\sin^2\theta} = a_\theta^2 \tag{1.220}$$

$$\left(\frac{dW_r}{dr}\right)^2 + \frac{a_\theta^2}{r^2} = 2m\left[E - V(r)\right] \tag{1.221}$$

where a_φ and a_θ are constants. Remembering that W will be a generator of a canonical transformation $r, \theta, \varphi, p_r, p_\theta, p_\varphi \rightarrow w_r, w_\theta, w_\varphi, J_r, J_\theta, J_\varphi$, we have

$$p_\varphi = \frac{\partial W}{\partial \varphi} = a_\varphi \tag{1.222}$$

$$p_\theta = \frac{\partial W}{\partial \theta} = \sqrt{a_\theta^2 - \frac{a_\varphi^2}{\sin^2\theta}} \tag{1.223}$$

$$p_r = \frac{\partial W}{\partial r} = \sqrt{2m\left[E - V(r) - \frac{a_\theta^2}{r^2}\right]} \tag{1.224}$$

The total orbital angular momentum is given by

$$L = mr\sqrt{(r\dot\theta)^2 + (r\sin\theta\,\dot\varphi)^2}$$

$$= \sqrt{p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}} = a_\theta \tag{1.225}$$

while the Z-component is given by

$$L_Z = mr^2\sin^2\theta\,\dot\varphi = p_\varphi = a_\varphi . \tag{1.226}$$

We shall now consider only the case of bound states of the particle which moves in the central force field. In that case $E < 0$ and we shall assume that $V(r)$ is everywhere negative but goes to zero,

monotically as $r \to \infty$. In that case the orbit is such that r oscillates between the two values, r_{min} and r_{max}, which make the radicand on the right of (1.224) vanish. Furthermore, it is evident from (1.223) that $|a_\varphi|$ must be less than a_θ and that θ oscillates between the values $\frac{\pi}{2} - \alpha$ and $\frac{\pi}{2} + \alpha$ where

$$\cos\alpha = \frac{|a_\varphi|}{a_\theta}, \tag{1.227}$$

α is the angle of inclination of the orbit.

The action variables are now given by

$$J_\varphi = \frac{1}{2\pi}\oint p_\varphi \, d\varphi = a_\varphi \tag{1.228}$$

$$J_\theta = \frac{1}{2\pi}\oint p_\theta \, d\theta$$

$$= \frac{1}{\pi}\int_{\pi/2-\alpha}^{\pi/2+\alpha} \sqrt{a_\theta^2 - \frac{a_\varphi^2}{\sin^2\theta}}\, d\theta \geq 0 \tag{1.229}$$

$$J_r = \frac{1}{2\pi}\oint p_r \, dr$$

$$= \frac{1}{\pi}\int_{r_{min}}^{r_{max}} \sqrt{2m[E - V(r)] - \frac{a_\theta^2}{r^2}}\, dr \geq 0 \tag{1.230}$$

The integral (1.229) can be evaluated by making the substitution $x = \cos\theta$ We then have

$$J_\theta = \frac{1}{\pi}\int_{\sin\alpha}^{-\sin\alpha} \sqrt{a_\theta^2 - \frac{a_\varphi^2}{1-x^2}}\, \frac{dx}{\sqrt{1-x^2}}$$

$$= \frac{1}{\pi}\int_{-\sin\alpha}^{\sin\alpha} \frac{\sqrt{(a_\theta^2 - a_\varphi^2) - a_\theta^2 x^2}}{1-x^2}\, dx \tag{1.231}$$

$$= \frac{a_\theta}{\pi}\int_{-\sin\alpha}^{\sin\alpha} \frac{\sqrt{\sin^2\alpha - x^2}}{1-x^2}\, dx$$

Consider the integral

$$I(a) = \int_{-a}^{a} \frac{\sqrt{a^2 - x^2}}{1 - x^2}\, dx \tag{1.232}$$

We have

$$I'(a) = 2a \int_0^a \frac{dx}{(1-x^2)\sqrt{a^2 - x^2}}$$

$$= 2a \left[\frac{1}{\sqrt{1-a^2}} \tan^{-1}\left(x \sqrt{\frac{1-a^2}{a^2 - x^2}} \right) \right]_0^a$$

$$= \frac{\pi a}{\sqrt{1-a^2}}\,, \tag{1.233}$$

and hence, remembering $I(0) = 0$,

$$I(a) = \pi \left[1 - \sqrt{1-a^2} \right]. \tag{1.234}$$

Thus, finally,

$$\gamma_\theta = a_\theta \left[1 - \cos \alpha \right] = a_\theta - |a_\varphi| \tag{1.235}$$

We may now write

$$J_r = \frac{1}{\pi} \int_{r_{min}}^{r_{max}} \sqrt{2m\left[E - V(r)\right] - \frac{\left(J_\theta + J_\varphi\right)^2}{r^2}}\, dr, \tag{1.231}$$

In principle, equation (1.231) may be solved for E in terms of J_r, J_θ, J_φ. It will be noted that E depends on J_θ and J_φ only through the combination $J_\theta + J_\varphi$. Therefore

$$\omega_\theta = |\omega_\varphi| = \frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial J_\varphi}\,, \tag{1.232}$$

and central force motion is seen to be degenerate.

We may therefore transform to a new set of angle and action variables for which one of the frequencies vanish. This is most conveniently done as follows :

$$J_n = J_r + J_\theta + |J_\varphi| \tag{1.233}$$

$$J_\ell = J_\theta + |J_\varphi| \tag{1.234}$$

$$J_m = J_\varphi \tag{1.235}$$

and

$$\omega_n = \omega_r \tag{1.236}$$

$$\omega_\ell = -\omega_r + \omega_\theta \tag{1.237}$$

$$\omega_m = \mp \omega_\theta + \omega_\varphi \tag{1.238}$$

The matrix of the transformation is

$\mp$ (+ sign when $J \gtrless 0$; − sign when $J \lessgtr 0$)

$$\tau = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & \mp 1 & 1 \end{pmatrix} \quad \tau^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \pm 1 & \pm 1 & 1 \end{pmatrix} \tag{1.239}$$

It has integral elements and determinent unity.

It is evident that

$$J_\ell = a_\theta = L , \tag{1.240}$$

$$J_m = a_\varphi = L_Z , \tag{1.241}$$

and

$$\omega_m = \frac{\partial E}{\partial J_m} = 0 , \tag{1.242}$$

J_ℓ and J_n may be taken positive, and we have the restrictions

$$-J_\ell \leq J_m \leq J_\ell \quad , \quad J_\ell \leq J_n . \tag{1.243}$$

In order to analyze the orbit it is sufficient to consider motion in a plane. The Lagrangian function will be

$$L = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) - V(r) \tag{1.244}$$

and this will lead to a space-action of the form

$$W = W_r(r) + W_\theta(\theta) \tag{1.245}$$

where

$$p_\theta = \frac{dW_\theta}{d\theta} = J_\theta \tag{1.246}$$

and

$$p_r = \frac{dW_r}{dr} = \sqrt{2m[E - V(r)] - \frac{J_\theta^2}{r^2}} . \tag{1.247}$$

We may then write

$$w_r = \frac{\partial W}{\partial J_r} = \frac{\partial W_r}{\partial J_r} = \frac{\partial}{\partial J_r} \int p_r \, dr = \int \frac{\partial p_r}{\partial J_r} \, dr$$

$$= \int \frac{m \omega_r}{\sqrt{2m[E - V(r) - J_\theta^2/r^2}} \, dr \tag{1.248}$$

where $\omega_r = \partial E / \partial J_r$. Since $w_r = \omega_r t$, we have

$$dt = \frac{dw_r}{\omega_r} = \frac{m \, dr}{\sqrt{2m[E - V(r)] - J_\theta^2/r^2}} \tag{1.249}$$

But since

$$J_\theta = p_\theta = m r^2 \frac{d\theta}{dt} \tag{1.250}$$

we have

$$\frac{d\theta}{dr} = \frac{d\theta}{dt}\frac{dt}{dr} = \frac{J_\theta}{m r^2}\frac{dt}{dr}$$

$$= \frac{J_\theta}{r^2\sqrt{2m[E - V(r)] - J_\theta^2/r^2}} \tag{1.251}$$

Integration of (1.251) gives the equation of the orbit.

In the case of Kepler motion we have

$$V(r) = -\frac{Ze^2}{r} \tag{1.252}$$

and hence

$$J_r = \frac{1}{\pi}\int_{r_{min}}^{r_{max}} \sqrt{2mE + \frac{2mZe^2}{r} - \frac{J_e^2}{r^2}}\, dr. \tag{1.253}$$

This integral may be evaluated as follows. Set

$$2mE = -A \quad , \quad mZe^2 = B \quad , \quad J_e^2 = C \tag{1.254}$$

Then

$$J_r = \frac{1}{\pi}\int_{r_{min}}^{r_{max}} \frac{1}{r}\sqrt{-Ar^2 + 2Br + C}\, dr$$

$$= \frac{\sqrt{A}}{\pi}\int_{-a}^{a} \frac{1}{x + B/A}\sqrt{a^2 - x^2}\, dx \tag{1.255}$$

where

$$x = r - \frac{B}{A} \quad , \quad a^2 = \frac{B^2 - AC}{A^2} \tag{1.256}$$

and

$$\left.\begin{aligned} r_{min} &= \frac{B}{A} - \frac{1}{A}\sqrt{B^2 - AC} \\ r_{max} &= \frac{B}{A} + \frac{1}{A}\sqrt{B^2 - AC} \end{aligned}\right\} \tag{1.257}$$

Differentiating J_2 with respect to a we obtain

$$\frac{dJ_2}{da} = \frac{a\sqrt{A}}{\pi}\int_{-a}^{a} \frac{dx}{(x+\frac{B}{A})\sqrt{a^2-x^2}}$$

$$= \frac{a\sqrt{A}}{\pi\sqrt{B^2/A^2-a^2}}\left[\sin^{-1}\frac{a^2+\frac{B}{A}x}{a(x+\frac{B}{A})}\right]_{-a}^{a}$$

$$= \frac{a\sqrt{A}}{\sqrt{B^2/A^2-a^2}} \qquad (1.258)$$

Integrating, and remembering that $J_2 = 0$ when $a = 0$, we have

$$J_2 = \frac{B}{\sqrt{A}} - \sqrt{A}\sqrt{\frac{B^2}{A^2}-a^2}$$

$$= \frac{B}{\sqrt{A}} - \sqrt{C}$$

$$= \frac{m Z^2 e^2}{\sqrt{-2mE}} - J_\ell \qquad (1.259)$$

and hence finally

$$E = -\frac{m Z^2 e^4}{2(J_r+J_\ell)^2} = -\frac{m Z^2 e^4}{2\hbar^2 n^2} \qquad (1.260)$$

Since E depends on only the one action variable J_n, Kepler motion is seen to be 2-fold degenerate.

$$\omega_m = \frac{\partial E}{\partial J_m} = 0 \quad , \quad \omega_\ell = \frac{\partial E}{\partial J_\ell} = 0 \qquad (1.261)$$

The orbit is a self-closing curve, and the frequency of rotation in this orbit is

$$\omega_n = \frac{\partial E}{\partial J_n} = \frac{m Z^2 e^4}{J_n^3} \, . \qquad (1.262)$$

Problem XXI : Find the action variables for a relativistic particle moving in a central force field, for which the Lagrangian function is

$$L = -mc^2\sqrt{1-\frac{\dot{r}^2}{c^2}} - V(r)$$

leading to the Hamiltonian function:

$$H = c\sqrt{p^2+m^2c^2} + V(r)$$

Show that, as in the non-relativistic case,

$$J_\varphi = L_z \quad , \quad J_\theta = L - L_z$$

where L is the magnitude of the total orbital angular momentum and L_z is the z-component. Show also that

$$J_r = \frac{1}{\pi}\int_{r_{min}}^{r_{max}} \sqrt{\frac{1}{c^2}[E-V(r)]^2 - m^2c^2 - \frac{(J_\theta+|J_\varphi|)^2}{r^2}}\,dr$$

so that the motion is 1-fold degenerate.

Show finally that for Kepler motion, with $V(r) = -\frac{Ze^2}{r}$

$$E = \frac{mc^2}{\sqrt{1+\frac{Z^2e^4}{c^2(J_n-J_\theta+\sqrt{J_\theta^2-Z^2e^4/c^2})^2}}}$$

where $J_\theta = J_\theta + |J_\varphi|$, $J_n = J_r + J_\theta + J_\varphi|$, and hence that we must have $J_\theta > \frac{Ze^2}{c}$ for a real bound state. Show that for $\frac{Ze^2}{c} \ll J_n, J_\theta$ this equation for the energy may be approximated in the form

$$E \approx mc^2\left\{1 - \frac{1}{2}\left(\frac{Ze^2}{c}\right)^2\frac{1}{J_n^2} - \frac{1}{2}\left(\frac{Ze^2}{c}\right)^4\frac{1}{J_n^4}\left(\frac{J_n}{J_\theta}-\frac{3}{4}\right)\right\}.$$

Bibliography

M. Born, The Mechanics of the Atom, London, G. Bell and Sons, Ltd, 1927

H. Goldstein, Classical Mechanics, Addison-Wesley Press, Inc., 1950

E.T. Whittaker, Analytical Dynamics, Cambridge University Press, 1937

2. CLASSICAL PERTURBATION THEORY.

Consider a system, with Hamiltonian function H_o which has multiply-periodic motions. Consider a second system which is described by a Hamiltonian function

$$H = H_o + H_1 \tag{2.1}$$

which differs from that of the first system by an amount H_1. If H_1 is sufficiently small the second system should possess motions which approximate the multiply periodic motions of the first system. It is the purpose of this section to derive approximation procedures for obtaining these "perturbed' motions.

We consider first the case in which H_1, as well as H_o, is not explicitly dependent upon the time. The motions of the perturbed system may then be themselves regarded as multiply periodic, at least over appreciable periods of time, and should be describable in terms of angle and action variables w^i, J_i. Our problem will be to obtain these angle and action variables in terms of the angle and action variables w_o^i, J_{oi} of the unperturbed system. We shall do this by finding a generating function $W(w_o, J)$ which leads from the variables w_o^i, J_{oi} to the variables w^i, J_i by equations of the form

$$J_{oi} = \frac{\partial W}{\partial w_o^i}, \qquad w^i = \frac{\partial W}{\partial J_i} \tag{2.2}$$

(Cf. Eq. (1.153)).

The theory of this section finds its application in cases in which the Hamiltonian H_o of the unperturbed system is simple and where the solutions of the equations for the unperturbed motion are well-known. Accordingly we shall assume that the angle and action variables w_o^i; J_{oi} of the unperturbed system are given as known functions of whatever canonical

variables q_i, p_i the system may initially have been described in terms of. H_0 and H_1 may therefore be taken as known functions of the w_{0i} and J_{0i}. H_0 is, of course, dependent only upon the J_{0i}, being simply the energy function E_0 (J_0). Thus we write

$$H\ (w_0,\ J_0) = H_0\ (J_0) + H_1\ (w_0,\ J_0) \tag{2.3}$$

Although the w_{0i} and J_{0i} are canonical variables they are not, in general, the angle and action variables of the perturbed system. In order to find these new variables w_i, J_i, we must first find the generating function of Eq. (2.2) by solving the Hamilton-Jacobi equation

$$H\left(w_0, \frac{\partial W}{\partial w_0}\right) = H_0\left(\frac{\partial W}{\partial w_0}\right) + H_1\left(w_0, \frac{\partial W}{\partial w_0}\right) = E \tag{2.4}$$

We solve this equation by a method of successive approximations, writing

$$W = W_0 + W_1 + W_2 + \ldots, \tag{2.5}$$

$$E = E_0 + E_1 + E_2 + \ldots. \tag{2.6}$$

The subscripts on the terms of these series have the following significance. We imagine that the perturbing part of the Hamiltonian function, namely H_1, is proportional to a small parameter α. The subscripts then denote the order in α of the given term, and the terms are to be evaluated by expanding the left and right sides of equation (2.4) in powers of α and equating terms of equal order. To zeroth order the transformation $w_{0i}, J_{0i} \rightarrow w_i, J_i$ must be simply the identity transformation. Hence

$$W_0\ (w_0,\ J) = w_{0i} J_i \tag{2.7}$$

and $$E_0(\ J\ ) = H_0\ (J), \tag{2.8}$$

E_0 is, as before, just the energy function of the unperturbed system. Now, however, we are regarding it as a function of the J's rather than

of the J_0's.

We now make an important assumption, namely, that the particular unperturbed motion under consideration, which serves as a point of departure for our approximation procedure, shall be non-degenerate, not even accidentally degenerate. This means that there exist no integers τ_i, except zeros, such that $\tau_i \omega_0^i = 0$, where $\omega_0^i = \partial E_0(J_0) / \partial J_{0i}$. This assumption will be discussed in greater detail presently.

Since the ω_0^i, J_{0i} and ω^i, J_i are separately the angle and action variables of two different systems, the initial canonical variables q^i, p_i must be periodic (with period 2π) separately in the ω_0^i and in the ω^i. This means that, apart from an arbitrary integral linear transformation with determinent ∓ 1 of the angle variables among themselves, we must have

$$\omega^i = \omega_0^i + f^i(\omega_0) \tag{2.9}$$

where the $f^i(\omega_0)$ are certain functions which are periodic in the ω_0's with period 2π. We should now remember that one final condition must be satisfied in order that the ω^i be true angle variables, namely that the function

$$W = W - \omega^i J_i \tag{2.10}$$

be periodic in the ω^i's with period 2π. In view of (2.9) this means that W, considered as a function of the ω_0^i's, must also be periodic in these latter variables with the period 2π. But

$$\begin{aligned} W &= W - \omega_0^i J_i - f^i J_i \\ &= W_1 + W_2 + \cdots - f^i J_i \end{aligned} \tag{2.11}$$

Hence W_1, W_2, . . . must all be periodic in the ω_0^i's with period 2π. Hence we may write

$$W_r(\omega_0, J) = \sum_{\tau}{}' W_{r\tau}(J)\, e^{i\tau\omega_0} \quad \text{for all } r \geq 1 \tag{2.12}$$

where the subscript τ is short for $\tau_1 \ldots \tau_n$, the notation $\tau\omega_0$ is an abbreviation for $\tau_i \omega_0^i$ and the summation is to be carried out over all positive and negative integral values of the τ_i, including zero. From the reality condition on the W_r we must have

$$W^*_{r\tau} = W_{r-\tau} \quad \text{for all } r \geq 1 \text{ and } \tau \tag{2.13}$$

Without loss of generality we may also choose

$$W_{r0} = 0 \quad \text{for all } r \geq 1, \tag{2.14}$$

for the only effect which a choice of non-vanishing values for the constant terms in the series (2.12) has is to shift the point of origin of the angle variables ω^i.

Since the initial canonical variables ρ^i, p_i are periodic in the ω_0's, and since the perturbing term, H_1, in the Hamiltonian is originally expressed in terms of the ρ^i, p_i, it is evident that H_1 is itself periodic in the ω_0's with period 2π. Thus we may write

$$H_1(\omega_0, J_0) = \sum_{\tau} H_{1\tau}(J_0)\, e^{i\tau\omega_0} \tag{2.15}$$

$$\text{or } H_1(\omega_0, J) = \sum_{\tau} H_{1\tau}(J)\, e^{i\tau\omega_0}, \tag{2.16}$$

together with the reality condition

$$H^*_{1\tau} = H_{1-\tau} \quad \text{for all } \tau. \tag{2.17}$$

Let us now expand the Hamilton-Jacobi equation (2.4). We obtain, using (2.7),

$$E_0(J) + E_1(J) + E_2(J) + \ldots$$

$$= H_0\left(J + \frac{\partial W_1}{\partial w_0} + \frac{\partial W_2}{\partial w_0} + \dots\right) + H_1\left(w_0, J + \frac{\partial W_1}{\partial w_0} + \frac{\partial W_2}{\partial w_0} + \dots\right)$$

$$= H_0(J) + \frac{\partial H_0(J)}{\partial J_i}\left(\frac{\partial W_1(w_0, J)}{\partial w_0^i} + \frac{\partial W_2(w_0, J)}{\partial w_0^i} + \dots\right)$$

$$+ \frac{1}{2!}\frac{\partial^2 H_0(J)}{\partial J_i \partial J_j}\left(\frac{\partial W_1(w_0, J)}{\partial w_0^i} + \frac{\partial W_2(w_0, J)}{\partial w_0^i} + \dots\right)\left(\frac{\partial W_1(w_0, J)}{\partial w_0^j} + \frac{\partial W_2(w_0, J)}{\partial w_0^j} + \dots\right)$$

$$+ \dots\dots$$

$$+ H_1(w_0, J) + \frac{\partial H_1(w_0, J)}{\partial J_i}\left(\frac{\partial W_1(w_0, J)}{\partial w_0^i} + \frac{\partial W_2(w_0, J)}{\partial w_0^i} + \dots\right)$$

$$+ \dots \qquad (2.18)$$

Equating terms of equal order in α we obtain the following set of equations to serve as the basis of our method of successive approximations :

$$E_0(J) = H_0(J) \qquad (2.19)$$

$$E_1(J) = H_1(w_0, J) + \frac{\partial H_0(J)}{\partial J_i}\frac{\partial W_1(w_0, J)}{\partial w_0^i} \qquad (2.20)$$

$$E_2(J) = \frac{\partial H_1(w_0, J)}{\partial J_i}\frac{\partial W_1(w_0, J)}{\partial w_0^i} + \frac{1}{2}\frac{\partial^2 H_0(J)}{\partial J_i \partial J_j}\frac{\partial W_1(w_0, J)}{\partial w_0^i}\frac{\partial W_1(w_0, J)}{\partial w_0^j} + \frac{\partial H_0(J)}{\partial J_i}\frac{\partial W_2(w_0, J)}{\partial w_0^i} \qquad (2.21)$$

. . . . . .

Equations (2.20), (2.21) etc.. connect quantities on the left which are functions of the J's alone to quantities on the right which are functions

of both the ω_0's and the J's. The quantities on the right are, however, periodic in the ω_0's. Hence the quantities on the left are given simply by the sum of the constant terms in the periodic quantities on the right. The sum of the oscillating terms must vanish. The constant terms are obtained by averaging over the ω_0^{-1}s. This is equivalent to taking a time average with respect to the unperturbed motion. Denoting such averages by a bar, we have

$$E_1 = \overline{H_1} = H_{10} \tag{2.22}$$

$$E_2 = \overline{\frac{\partial H_1}{\partial J_i}\frac{\partial W_1}{\partial \omega_0^i}} + \frac{1}{2}\overline{\frac{\partial^2 H_0}{\partial J_i \partial J_j}\frac{\partial W_1}{\partial \omega_0^i}\frac{\partial W_1}{\partial \omega_0^j}}$$

$$= -\sum_\tau{}' i\tau_i \frac{\partial H_{1\tau}}{\partial J_i} W_{1-\tau} + \frac{1}{2}\frac{\partial^2 H_0}{\partial J_i \partial J_j}\sum_\tau{}' \tau_i \tau_j W_{1\tau} W_{1-\tau} \tag{2.23}$$

. . . . . etc.

From (2.22) it is seen that the energy of the perturbed motion is, to first approximation, equal to the energy of the unperturbed motion plus the time average of H_1 taken over the unperturbed motion. To obtain the energy in the second approximation we must obtain explicit expressions for the amplitudes $W_{1\tau}$. This we do as follows. Subtracting (2.22) from (2.20) we obtain

$$0 = \sum_\tau{}' H_{1\tau} e^{i\tau\omega_0} + \omega_0^i \sum_\tau{}' i\tau_i W_{1\tau} e^{i\tau\omega_0} \tag{2.24}$$

where the prime indicates that the constant term ($\tau = 0$) is to be omitted from the summation and where $\omega_0^i = \omega_0^i(J) = \partial H_0(J)/\partial J_i$. We thus find

$$W_{1\tau} = \frac{iH_{1\tau}}{\tau\omega_0} \qquad \text{for } \tau \neq 0 \tag{2.25}$$

Since the unperturbed motion is non-degenerate the denominator on the right

of (2.25) never vanishes. We have finally

$$E_2 = -\sum_{\tau}{}' \tau_i \frac{\partial H_{1\tau}}{\partial J_i} \frac{H_{1-\tau}}{\tau\omega_0} + \frac{1}{2}\frac{\partial \omega_0^i}{\partial J_i} \sum_{\tau}{}'' \tau_i \tau_j \frac{H_{1\tau}}{\tau\omega_0} \frac{H_{1-\tau}}{\tau\omega_0}$$

$$= -\frac{1}{2}\sum_{\tau}{}' \tau_i \left\{ \frac{\partial H_{1\tau}}{\partial J_i} \frac{H_{1-\tau}}{\tau\omega_0} + \frac{\partial H_{1-\tau}}{\partial J_i} \frac{H_{1\tau}}{\tau\omega_0} + H_{1\tau} H_{1-\tau} \frac{\partial}{\partial J_i}\left(\frac{1}{\tau\omega_0}\right) \right\}$$

$$= -\sum_{\tau\omega_0 > 0} \tau_i \frac{\partial}{\partial J_i}\left(\frac{|H_{1\tau}|^2}{\tau\omega_0}\right) \qquad (2.26)$$

In principle we may continue this formal procedure indefinitely, obtaining series of the form

$$E = H_0 + H_{10} - \sum_{\tau\omega_0 > 0} \tau_i \frac{\partial}{\partial J_i}\left(\frac{|H_{1\tau}|^2}{\tau\omega_0}\right) + \ldots \qquad (2.27)$$

$$W = \omega_0^i J_i + \sum_{\tau}{}' \frac{i H_{1\tau}}{\tau\omega_0} e^{i\tau\omega_0} + \ldots \qquad (2.28)$$

The fundamental problem of determining the convergence of these series has never been completely resolved. Indeed, entirely aside from the question of the convergence of these series as a whole, there is the question of the convergence of the individual terms, which are themselves infinite series. In the case of degeneracy in the unperturbed motion it is quite clear that even the terms diverge, owing to the presence of the denominators $\tau\omega_0$ (which appear generally in all the higher terms). Even in the case of non-degeneracy in the unperturbed motion these denominators become arbitrarily small by choosing the integers τ_i sufficiently large. The question then becomes one of determining whether or not the numerators $H_{1\tau}$, $|H_{1\tau}|^2$, etc. become sufficiently small with large τ_i to neutralize the effect of the small denominators. It has actually been possible to exhibit perfectly respectable non-degenerate dynamical systems for which the series do converge. (See, for example; Whittaker, Analytical Dynamics, §§ 197; 198)

The case of initial degeneracy requires a modified treatment

into which we shall not go here. (For details the reader is referred to Born's Mechanics of the Atom §§ 43-47.) A few remarks should be made, however, concerning accidental initial degeneracy. For the vast majority of multiply-periodic systems, each orbit which is non-degenerate lies infinitely close to an orbit which is accidentally degenerate. (See H. Bruns Astronomische Nachrichtung, CIX, p. 215 (1884).) This means that the series (2.27, 28) represent completely discontinuous functions of the action variables J_i. This can readily be illustrated by the example of a two-dimensional particle in a square box of side ℓ. From the results of Problem XX it is evident that the energy of the particle is given by

$$E = \frac{\pi^2}{2m\ell^2}\left(J_1^2 + J_2^2\right), \tag{2.29}$$

and hence

$$\omega_1 = \frac{\partial E}{\partial J_1} = \frac{\pi^2}{m\ell^2} J_1 \quad , \quad \omega_2 = \frac{\partial E}{\partial J_2} = \frac{\pi^2}{m\ell^2} J_2 \tag{2.30}$$

The motion is accidentally degenerate or non-degenerate according as the ratio

$$\frac{\omega_1}{\omega_2} = \frac{J_1}{J_2} \tag{2.31}$$

is rational or irrational. As J_1 (or J_2) varies continuously the orbit varies continuously, but its degeneracy or non-degeneracy varies in a completely discontinuous fashion.

It is, of course, true that the non-degenerate orbits are infinitely more numerous than the accidentally degenerate ones, and this has formed the basis of classical ergodic theory. It also forms the basis of the following remark : even for initially degenerate motion it may happen that the series (2.27 , 28) represent a good approximation to the perturbed motion over long (though not infinite) periods of time, provided the series (particularly the individual term series) are terminated after the first few terms (in any case before the denominators $\tau\omega_3$ vanish).

The accuracy of the resulting representation depends upon the rapidity with wich the numerators $H_{1\tau}, |H_{1\tau}|^2$, etc. tend to zero with increasing τ_i values. (See H. Poincaré, Méthodes Nouvelles de la Mécanique Céleste, Paris, 1892-99, vol. I, chap. V; vol. II, chap. VII.)

From (2.2), (2.7) and (2.12) we may write

$$J_{oi} = J_i + \sum_{r,\tau} i\tau_i W_{r\tau}(J) e^{i\tau\omega_o} \qquad (2.32)$$

From (2.14) it is therefore evident that the values of the new action variables are simply equal to the time average of the old action variables, averaged over the motion of the unperturbed system.

$$J_i = \overline{J_{oi}} \qquad (2.33)$$

We also have

$$\omega^i = \omega_o{}^i + \sum_{r,\tau} \frac{\partial W_{r\tau}(J)}{\partial J_i} e^{i\tau\omega_o} \qquad (2.34)$$

which identifies for us the functions f^i of (2.9).

In many cases of importance the perturbing term, H_1, in the Hamiltonian function is explicitly dependent on the time. When this situation occurs we may still construct a perturbation theory in a manner analogous to the preceding theory. Since the Hamiltonian function is no longer time independent, the perturbed system will not undergo truly multiply periodic motion. We shall continue, however, to seek for a canonical transformation of the form (2.2) which transforms the Hamiltonian function into a function which is explicitly dependent on only one of the sets of canonical variables, the J's, together, now, with the time. The generating function W of the transformation will now also be explicitly dependent on the time. Hence, referring to (1.19), we see that the generalization of the Hamilton-Jacobi equation (2.4) for the present case has the form

$$H_o\left(\frac{\partial W}{\partial \omega_o}\right) + H_1\left(\omega_o, \frac{\partial W}{\partial \omega_o}, t\right) + \frac{\partial W}{\partial t} = E(J, t) \qquad (2.35)$$

We again solve the Hamilton-Jacobi equation by successive approximations, making use of expansions of the form (2.5. 6), and we again make the assumption of initial non-degeneracy. As previously we may write

$$H_1(\omega_0, J, t) = \sum_\tau H_{1\tau}(J, t) e^{i\tau\omega_0}, \tag{2.36}$$

and, in analogy with (2.12), we make the ansatz

$$W_r(\omega_0, J, t) = \sum_\tau W_{r\tau}(J, t) e^{i\tau\omega_0} \quad \text{for } r \geq 1 \tag{2.37}$$

$$W_{r_0} = 0 \quad \text{for all } r \geq 1. \tag{2.38}$$

We have again, of course,

$$W_0 = \omega_0{}^i J_i. \tag{2.39}$$

The expansion of the Hamilton-Jacobi equation has the form

$$\begin{aligned} & E_0(J) + E_1(J, t) + \dots \\ & = H_0(J) + \frac{\partial H_0(J)}{\partial J_i}\left(\frac{\partial W_1(\omega_0, J, t)}{\partial \omega_0{}^i} + \dots\right) + \dots \\ & + H_1(\omega_0, J, t) + \dots \\ & + \frac{\partial W_1(\omega_0, J, t)}{\partial t} + \dots, \end{aligned} \tag{2.40}$$

leading to

$$E_0(J) = H_0(J) \tag{2.41}$$

$$E_1(J, t) = H_1(\omega_0, J, t) + \frac{\partial W_1(\omega_0, J, t)}{\partial t} + \frac{\partial H_0(J)}{\partial J_i}\frac{\partial W_1(\omega_0, J, t)}{\partial \omega_0{}^i} \tag{2.42}$$

. . . . . .

Averaging Eq. (2.42) over the ω_0's, we obtain

$$E_1 = H_{10} \tag{2.43}$$

and

$$0 = \sum_\tau' \left[H_{1\tau} + \frac{\partial W_{1\tau}}{\partial t} + i\tau_i \omega_0 i W_{1\tau} \right] e^{i\tau\omega_0} . \tag{2.44}$$

If we now imagine the perturbing term H_1 to be "switched on" at time $-\frac{T}{2}$, so that W_1 vanishes at $t = -\frac{T}{2}$, the solution of (2.44) is given by

$$W_{1\tau}(\gamma, t) = -\int_{T/2}^{t} H_{1\tau}(\gamma, t') e^{i\tau\omega_0(\gamma)(t'-t)} dt' \qquad \text{for } \tau \neq 0 \tag{2.45}$$

Equation (2.45) may readily be seen to reduce to (2.25) when H_1 is not explicitly dependent on the time. For, using (2.13) and (2.16) of the Appendix, we then have, taking $T \to \infty$,

$$\begin{aligned} W_{1\tau} &= -H_{1\tau} \int_{-\infty}^{t} e^{i\tau\omega_0(t'-t)} dt' \\ &= -H_{1\tau} \int_{-\infty}^{0} e^{i\tau\omega_0 t''} dt'' \qquad t'' = t'-t \\ &= -2\pi H_{1\tau} \delta_-(\tau\omega_0) \\ &= -2\pi H_{1\tau} \left[\frac{1}{2}\delta(\tau\omega_0) - \frac{i}{2\pi}\frac{1}{\tau\omega_0} \right] \\ &= \frac{iH_{1\tau}}{\tau\omega_0} \qquad \text{for } \tau \neq 0 \end{aligned} \tag{2.46}$$

We must have $\delta(\tau\omega_0) = 0$ since, owing to the assumed non-degeneracy

of the unperturbed motion, $\tau \neq 0$ implies $\tau\omega_0 \neq 0$.

The following relations will be useful to us :

$$\omega^i = \frac{\partial W}{\partial J_i} = \omega_0^i + \sum_\tau \frac{\partial W_{1\tau}}{\partial J_i} e^{i\tau\omega_0} + \ldots \qquad (2.47)$$

$$J_{0i} = \frac{\partial W}{\partial \omega_0^i} = J_i + \sum_\tau i\tau_i W_{1\tau} e^{i\tau\omega_0} + \ldots \qquad (2.48)$$

$$\dot{\omega}_0^i = \frac{\partial E}{\partial J_i} = \frac{\partial E_0}{\partial J_i} + \frac{\partial E_1}{\partial J_i} + \ldots = \omega_0^i + \frac{\partial H_0}{\partial J_i} + \ldots \qquad (2.49)$$

$$\dot{J}_i = -\frac{\partial E}{\partial \omega^i} = 0 \qquad (2.50)$$

We shall be interested in the actual value of the perturbing function H_1 at any instant of time. This is given, correct to second order, by

$$H_1(\omega_0, J_0, t) = H_1(\omega, J, t) - \frac{\partial H_1(\omega, J, t)}{\partial \omega_i} \frac{\partial W_1(\omega, J, t)}{\partial J_i} + \frac{\partial H_1(\omega, J, t)}{\partial J_i} \frac{\partial W_1(\omega, J, t)}{\partial \omega^i} + \ldots$$

$$= H_1 - (H_1, W_1) + \ldots$$

$$= \sum_\tau H_{1\tau} e^{i\tau\omega} - \sum_{\tau,\tau'} \left\{ i\tau_i H_{1\tau} \frac{\partial W_{1\tau'}}{\partial J_i} - i\tau_i' \frac{\partial H_{1\tau}}{\partial J_i} W_{1\tau'} \right\} e^{i(\tau+\tau')\omega} + \ldots \qquad (2.51)$$

where, in the second line, the quantities appearing are to be regarded as functions of the new canonical variables ω^i, J_i, the arguments ω_0^i being simply replaced by ω^i's. The variation in H_1 due to its explicit dependence on the time is also of interest. We have

$$\frac{\partial H_1(\omega_0, J_0, t)}{\partial t} = \left[\frac{\partial H_1}{\partial t} - \left(\frac{\partial H_1}{\partial t}, W_1 \right) + \ldots \right](\omega, J, t)$$

$$= \sum_\tau \frac{\partial H_{1\tau}}{\partial t} e^{i\tau\omega} - \sum_{\tau,\tau'} \left\{ i\tau_i \frac{\partial H_{1\tau}}{\partial t} \frac{\partial W_{1\tau'}}{\partial J_i} - i\tau_i' \frac{\partial^2 H_{1\tau}}{\partial t\, \partial J_i} W_{1\tau'} \right\} e^{i(\tau+\tau')\omega} + \ldots \qquad (2.52)$$

We shall now restrict our attention to systems for which the perturbing term in the Hamiltonian function is independent of the momenta p_i when expressed in terms of the original canonical variables q_i , p_i . That is

$$H_1(\omega_0, J_0, t) = H_1(q, t) = -L_1(q, t) \tag{2.53}$$

where L_1 is the perturbing term in the original Lagrangian function. The term L_1 describes the action of an external generalized force of amount

$$F_i^{ext} = \frac{\partial L_1}{\partial \dot{q}^i} = -\frac{\partial H_1}{\partial q^i} \tag{2.54}$$

The rate at which energy is imparted to the system through the action of this external force is given by

$$\frac{dE^{abs}}{dt} = F_i^{ext}\,\dot{q}^i = -\frac{\partial H_1}{\partial q^i}\,\dot{q}^i = -\frac{dH_1}{dt} + \frac{\partial H_1}{\partial t} \tag{2.55}$$

But, referring, to (2.51), (2.52) and (2.49), we have

$$\frac{dH_1}{dt} = \sum_\tau i\tau_i\left(\omega_0^i + \frac{\partial H_{10}}{\partial J_i}\right) H_{1\tau}\, e^{i\tau\omega} - \sum_{\tau,\tau'} i(\tau+\tau')\omega_0 \left\{ i\tau_i H_{1\tau}\frac{\partial W_{1\tau'}}{\partial J_i} - i\tau_i' \frac{\partial H_{1\tau}}{\partial J_i} W_{1\tau'} \right\} e^{i(\tau+\tau')\omega}$$

$$- \sum_{\tau,\tau'} \left\{ i\tau_i H_{1\tau} \frac{\partial^2 W_{1\tau'}}{\partial t\,\partial J_i} - i\tau_i' \frac{\partial H_{1\tau}}{\partial J_i}\frac{\partial W_{1\tau'}}{\partial t} \right\} e^{i(\tau+\tau')\omega} + \frac{\partial H_1}{\partial t} + \dots \tag{2.55}$$

Hence, making use of (2.44), we find, correct to the second order,

$$\frac{dE^{abs}}{dt} = -\sum_\tau i\tau_i\left(\omega_0^i + \frac{\partial H_{10}}{\partial J_i}\right) H_{1\tau}\, e^{i\tau\omega}$$

$$+ \sum_{\tau,\tau'} i(\tau+\tau')\omega_0 \left\{ i\tau_i H_{1\tau}\frac{\partial W_{1\tau'}}{\partial J_i} - i\tau_i' \frac{\partial H_{1\tau}}{\partial J_i} W_{1\tau'} \right\} e^{i(\tau+\tau')\omega}$$

$$+\sum_{\tau,\tau'}\Big\{-i\tau_i H_{1\tau}\frac{\partial H_{1\tau'}}{\partial J_i}+\tau_i H_{1\tau}\frac{\partial(\tau'\omega_0 W_{1\tau'})}{\partial J_i}$$

$$+i\tau_i\frac{\partial H_{1\tau}}{\partial J_i}H_{1\tau'}-\tau_i\frac{\partial H_{1\tau}}{\partial J_i}\tau'\omega_0 W_{1\tau'}\Big\}e^{i(\tau+\tau')\omega}+\ldots \tag{2.56}$$

Let us now apply this result to find the mean rate of absorption of energy by the members of an ensemble of identical systems. At the moment when the perturbing terms H_1 is switched on and the ω's begin to depart in value from the ω_0's, the ω's will be found in general to be randomly distributed in magnitude. That is, the periodic motions of the systems will have random phases. In order to get the mean rate of absorption of energy we must average over these phases. . This is equivalent to averaging over the ω's in (2.56). Thus we obtain

$$\frac{d\bar{E}^{abs}}{dt}=-\sum_{\tau}\Big\{i\tau_i H_{1\tau}\frac{\partial H_{1-\tau}}{\partial J_i}+\tau_i H_{1\tau}\frac{\partial(\tau\omega_0 W_{1-\tau})}{\partial J_i}$$

$$+i\tau_i\frac{\partial H_{1\tau}}{\partial J_i}H_{1-\tau}+\tau_i\frac{\partial H_{1\tau}}{\partial J_i}\tau\omega_0 W_{1-\tau}\Big\}+\ldots$$

$$=-\sum_{\tau}\Big\{i\tau_i\frac{\partial}{\partial J_i}(|H_{1\tau}|^2)+\tau_i\frac{\partial}{\partial J_i}(H_{1\tau}\tau\omega_0 W_{1-\tau})\Big\}+\ldots \tag{2.57}$$

Upon carrying out the summation we see that the first term in the last line of (2.57) vanishes owing to antisymmetry. We thus obtain for the total mean energy absorbed by the members of the ensemble during the time interval $-\frac{T}{2}$ to $\frac{T}{2}$,

$$\Delta E^{abs}=\int_{-T/2}^{T/2}\frac{d\bar{E}}{dt}dt=-\sum_{\tau}\int_{-T/2}^{T/2}\tau_i\frac{\partial}{\partial J_i}(H_{1\tau}\tau\omega_0 W_{1\tau})dt$$

$$=\sum_{\tau}\int_{-T/2}^{T/2}dt\int_{-T/2}^{t}dt'\,\tau_i\frac{\partial}{\partial J_i}\left[H_{1\tau}\tau\omega_0 H'_{1-\tau}e^{i\tau\omega_0(t-t')}\right]$$

$$= \frac{1}{2} \sum_{\tau}{}' \tau_i \frac{\partial}{\partial J_i} \left[\int_{-T/2}^{T/2} \int_{-T/2}^{T/2} H_{1\tau} \tau\omega_0 H'_{1-\tau} e^{i\tau\omega_0(t'-t)} dt' dt \right] \quad (2.58)$$

correct to second order. If the perturbation function is switched off again at $t = \frac{T}{2}$ then we may let the limits of integration above run from $-\infty$ to ∞, and (2.58) can be written in the more compact form

$$\Delta E^{abs} = \frac{1}{2} \sum_{\tau}{}' \tau_i \frac{\partial}{\partial J_i} (\tau\omega_0 k_{1\tau} k_{1-\tau})$$

$$= \sum_{\tau\omega_0 > 0} \tau_i \frac{\partial}{\partial J_i} (\tau\omega_0 |k_{1\tau}|^2) \quad (2.59)$$

where

$$k_{1\tau}(J) = \int_{-\infty}^{\infty} H_{1\tau}(J,t) e^{i\tau\omega_0(J)t} dt. \quad (2.60)$$

Expression (2.57) represents an average energy absorption rate taken over the members of a random ensemble of identical systems. It also represents a time average energy absorption rate for a single system, taken over a short period of time (of the order of the longest fundamental period of the unperturbed system), in the special case in which the $H_{1\tau}$'s vary very slowly with time. An example of such a case is that in which the perturbation function starts out from zero at some time in the remote past, is built up very slowly, and finally attains a constant functional form which is no longer explicitly dependent on the time. The perturbation function is then said to be switched on "adiabatically". We shall now consider this case in some detail.

Let us first apply Eq. (2.45) to the adiabatic case. Integrating by parts and taking $T \to \infty$, we obtain

$$W_{1\tau}(J,t) = -e^{i\tau\omega_0(J)t} \left\{ \left[H_{1\tau}(J,t') \frac{e^{i\tau\omega_0(J)t'}}{i\tau\omega_0(J)} \right]_{-\infty}^{t} - \int_{-\infty}^{t} H_{1\tau}(J,t') \frac{e^{i\tau\omega_0(J)t'}}{i\tau\omega_0(J)} dt' \right\}$$

$$= \frac{i H_{1\tau}(J,t)}{\tau\omega_0(J)} \quad (2.61)$$

where we have neglected the term in $\dot{H}_{1\tau}$ owing to the slowly varying nature of $H_{1\tau}$. Comparison with (2.25) shows us that the values of the $W_{1\tau}$'s and hence of W_1, are the same at any instant as if they had been calculated on the assumption that the $H_{1\tau}$'s are constant, having the values associated with the particular instant in question. The same will in general be true of W_2, W_3, . . . etc. The J's, generated from the w_o's and J_o's via Eqs. (2.47, 48), are therefore the true instantaneous action variables of the system. In particular, after, the $H_{1\tau}$'s have attained their final steady values, the motion of the system will be once again truly multiply periodic, and will be described by the action variables J_i. In the remote past, before the adiabatic building-up process has begun, the J's are equal to the original J_o's. Equation (2.50) shows that the J's remain constant in time. We have therefore the following theorem which expresses the so-called "adiabatic invariance" of the action variables :

Theorem : Although at the end of the adiabatic building-up process the system will be executing a different multiply periodic motion from that executed initially, the values of the action variables describing the motion in the two cases will be identical.

The energy absorbed by an adiabatically varying system is given by Eq. (2.59), with definition (2.60) replaced by

$$k_{1\tau}(J,t) = \int_{-\infty}^{t} H_{1\tau}(J,t)\, e^{i\tau\omega_0(J)t}\, dt$$

$$= \left[H_{1\tau}(J,t)\, \frac{e^{i\tau\omega_0(J)t}}{i\tau\omega_0(J)} \right]_{-\infty}^{t} - \int_{-\infty}^{t} \dot{H}_{1\tau}(J,t)\, \frac{e^{i\tau\omega_0(J)t}}{i\tau\omega_0(J)}\, dt$$

$$= H_{1\tau}(J,t)\, \frac{e^{i\tau\omega_0(J)t}}{i\tau\omega_0(J)} \qquad (2.62)$$

Inserting (2.62) into (2.59) we obtain

$$\Delta E^{abs}(t) = \sum_{\tau\omega_0>0} \tau_i \frac{\partial}{\partial \gamma_i}\left[\frac{|H_{1\tau}(\gamma,t)|^2}{\tau\omega_0(\gamma)}\right] \tag{2.63}$$

In particular, after the $H'_{1\tau}$s have attained their final steady values no more energy absorption takes place. On comparison with (2.26) one observes that the total energy absorbed is, to second order, equal to the negative of the second order correction to the total energy of the system produced by the perturbation H_1.

We shall now illustrate some of the preceding results by considering the example of the forced harmonic oscillator. If the external impressed force is $F(t)$ then the Lagrangian function is

$$L = \frac{m}{2}\dot{q}^2 - \frac{m\omega_0^2}{2}q^2 + qF(t) \tag{2.64}$$

yielding the equation of motion

$$m\ddot{q} + m\omega_0^2 q = F(t) \tag{2.65}$$

and the Hamiltonian function $H = H_0 + H_1$, where

$$H_0 = \frac{1}{2m}p^2 + \frac{m\omega_0^2}{2}q^2 \tag{2.66}$$

$$H_1 = -qF(t) \tag{2.67}$$

Transforming to the angle and action variables of the unperturbed system, we have, using (1.190),

$$q = \sqrt{\frac{J}{2m\omega_0}}\left(e^{i\omega_0} + e^{-i\omega_0}\right), \tag{2.68}$$

and hence

$$H_{10} = 0 \tag{2.69}$$

$$H_{11} = H_{1-1} = -\sqrt{\frac{I}{2m\omega_0}}\,F(t) \tag{2.70}$$

$$H_{1\tau} = 0 \quad \text{for} \quad \tau \neq \pm 1 \tag{2.71}$$

$$k_{11} = -\sqrt{\frac{I}{2m\omega_0}}\int_{-\infty}^{\infty} F(t)\, e^{i\omega_0 t}\, dt = -\sqrt{\frac{\pi I}{m\omega_0}}\,F^{T}(\omega_0) \tag{2.72}$$

$$= k_{1-1}^{*} \tag{2.72'}$$

$$k_{1\tau} = 0 \quad \text{for} \quad \tau \neq \mp 1. \tag{2.73}$$

Finally, therefore,

$$\Delta E^{abs} = \frac{\partial}{\partial J}\left(\omega_0 |k_{11}|^2\right) = \frac{\partial}{\partial J}\left(\frac{\pi I}{m}|F^{T}(\omega_0)|^2\right)$$

$$= \frac{\pi}{m}|F^{T}(\omega_0)|^2 \tag{2.74}$$

It turns out that equation (2.74) is not only correct to second order but exact. To show this we must solve the equation of motion (2.65). We can give an explicit solution provided we know the Green's functions of the operator $\frac{d^2}{dt^2} + \omega_0^2$. These Green's functions are quite easy to derive, using the mathematical methods of the Appendix. In order, however, to choose that Green's function which corresponds to appropriate boundary conditions we shall proceed from a more physical viewpoint.

Let us begin by replacing equation (2.65) by the equation

$$m\ddot{\rho} + 2m\varepsilon\dot{\rho} + m\omega_0^2\rho = F(t) \tag{2.75}$$

in which a "dissipative" term proportional to $\dot{\rho}$ and proportional to a positive infinitesimal ε has been added. If $F(t) = 0$ the solutions of this equation have the general form

$$\rho(t) = e^{-\epsilon t}\left(\rho_0 e^{i\omega_0 t} + \rho_0^* e^{-i\omega_0 t}\right) \qquad (2.76)$$

Free solutions of this form may always be superimposed upon the particular solutions of Eq. (2.75). It is to be observed, however, that, owing to the damping factor $e^{-\epsilon t}$, if the value of ρ is finite in the remote past when the force F is switched on then the "free" part of the solution will have completely died out by the time the moment of interest, t, is reached.

Let us now Fourier analyze ρ and F, writing

$$\rho(t) = \int_{-\infty}^{\infty} \chi(\omega) e^{i\omega t} d\omega, \text{ and } F(t) = \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega \qquad (2.77)$$

Since functions of the form (2.76) (i.e. damping functions) possess no Fourier transform, this procedure automatically eliminates any "free" part of the solution of (2.75). Equation (2.75) will now be satisfied if

$$\chi(\omega) = \frac{1}{m} \frac{f(\omega)}{\omega_0^2 - \omega^2 + 2i\epsilon\omega} \qquad (2.78)$$

But, using Eqs. (2.17, 18) of the Appendix, we may write

$$\frac{1}{\omega_0^2 - \omega^2 + 2i\epsilon\omega} = \frac{1}{(\omega_0^2 - \epsilon^2) - (\omega - i\epsilon)^2} = \frac{1}{\omega_0^2 - (\omega - i\epsilon)^2}$$

$$= \frac{1}{2\omega_0}\left[\frac{1}{\omega_0 + \omega - i\epsilon} + \frac{1}{\omega_0 - \omega + i\epsilon}\right]$$

$$= \frac{\pi i}{\omega_0}\left[\delta_-(\omega_0 + \omega) - \delta_+(\omega_0 - \omega)\right] \qquad (2.79)$$

Observing that

$$f(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(t) e^{-i\omega t} dt = \sqrt{2\pi}\, F^{T*}(\omega) \qquad (2.80)$$

and using Eqs. (2.15, 16) of the Appendix, we find, on substituting (2.78) into (2.77),

$$q(t) = \frac{\pi i}{m\omega_0}\int_{-\infty}^{\infty}\left[\delta_-(\omega_0+\omega) - \delta_+(\omega_0-\omega)\right] f(\omega) e^{i\omega t}\, d\omega$$

$$= \frac{\pi i}{(2\pi)^2 m\omega_0}\int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dt' \int_0^{\infty} dt'' F(t') \left\{ e^{-i(\omega_0+\omega)t''} e^{i\omega(t-t')} - e^{i(\omega_0-\omega)t''} e^{i\omega(t-t')} \right\}$$

$$= \frac{i}{2m\omega_0}\int_{-\infty}^{\infty} dt' \int_0^{\infty} dt'' F(t') \left(e^{-i\omega_0 t''} - e^{i\omega_0 t''}\right) \delta(t-t'-t'')$$

$$= \frac{1}{m\omega_0}\int_{-\infty}^{\infty} \frac{1}{2}\left(1 + \frac{t-t'}{|t-t'|}\right) \sin\omega_0(t-t') F(t')\, dt'$$

$$= \frac{1}{m\omega_0}\int_{-\infty}^{t} \sin\omega_0(t-t') F(t')\, dt'. \qquad (2.81)$$

We have shown, incidentally, that a possible Green's function for the operator $\frac{d^2}{dt^2} + \omega_0^2$ is

$$G^{ret}_{\frac{d^2}{dt^2}+\omega_0^2}(t) = \frac{1}{2\omega_0}\left(1 + \frac{t}{|t|}\right)\sin\omega_0 t \qquad (2.82)$$

This function may be called the "retarded" Green's function of the system (2.64). For, when used to solve the equation of motion (2.65), it has the effect that the position of the oscillating mass m at time t is determined solely by the "past history" of the system, i.e. by the values of the

impressed force F(t) at all previous times.

If we had inserted a negative dissipation term instead of a positive dissipation term in (2.75) then we would have arrived at the "advanced" Green's function :

$$G^{adv}_{\frac{d^2}{dt^2}+\omega_0^2}(t) = -\frac{1}{2\omega_0}\left(1-\frac{t}{|t|}\right)\sin\omega_0 t \qquad (2.83)$$

One half the sum of the Green's functions (2.82) and (2.83) is also a Green's function :

$$\overline{G}_{\frac{d^2}{dt^2}+\omega_0^2}(t) = \frac{1}{2\omega_0}\frac{t}{|t|}\sin\omega_0 t \qquad (2.84)$$

This function may be called the "mean-value" Green's function. It is the function which we would have obtained if we had left the infinitesimal ε out of equation (2.78) and had taken the principal value of the resulting Fourier integral.

The difference of functions (2.82) and (2.83), namely

$$G^{ret}_{\frac{d^2}{dt^2}+\omega_0^2}(t) - G^{adv}_{\frac{d^2}{dt^2}+\omega_0^2}(t) = \frac{1}{\omega_0}\sin\omega_0 t \qquad (2.85)$$

of course satisfies the free equation.

Using the solution (2.81) we may readily calculate the energy absorbed by the oscillator throughout the course of time. The omission of any "free" part from the solution (2.81) corresponds to the procedure of averaging over the phases of an ensemble of oscillators. We obtain

$$\Delta E^{abs} = \int_{-\infty}^{\infty} F(t)\,\dot{q}(t)\,dt$$

$$= \frac{1}{m}\int_{-\infty}^{\infty} dt \int_{-\infty}^{t} dt'\, F(t)\cos\omega_0(t-t')\,F(t')$$

$$= \frac{1}{2m}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(t)\cos\omega_0(t-t')\,F(t')\,dt\,dt'$$

$$= \frac{1}{4m} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(t) \left[e^{i\omega_0 (t-t')} + e^{-i\omega_0 (t'-t)} \right] F(t) \, dt \, dt'$$

$$= \frac{1}{2m} \left| \int_{-\infty}^{\infty} F(t) e^{-i\omega_0 t} \, dt \right|^2 = \frac{1}{2m} \left| \sqrt{2\pi} \, F^T(\omega_0) \right|^2$$

$$= \frac{\pi}{m} \left| F^T(\omega_0) \right|^2$$

which checks with (2.74).

Problem XXII : Show that the space-time action corresponding to the Lagrangian function (2.64) is given by

$$S(\rho, t / \rho', t')$$

$$= \frac{1}{\sin \omega_0 (t-t')} \left\{ \frac{m\omega_0}{2} \left[(\rho^2 + \rho'^2) \cos \omega_0 (t-t') - 2\rho\rho' \right] \right.$$

$$- \rho \int_{t'}^{t} \sin \omega_0 (t'-t'') F(t'') \, dt'' + \rho' \int_{t'}^{t} \sin \omega_0 (t-t'') F(t'') \, dt''$$

$$\left. + \frac{1}{m\omega_0} \int_{t'}^{t} dt'' \int_{t}^{t''} dt''' \sin \omega_0 (t'-t''') \sin \omega_0 (t-t'') F(t'') F(t''') \right\}$$

Show that this function satisfies the Hamilton-Jacobi equation for the system.

Bibliography

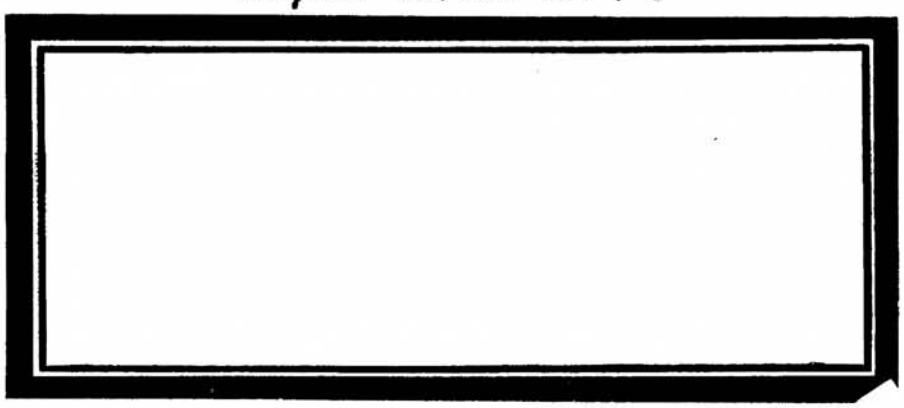
M. Born, The Mechanics of the Atom, London, G. Bell and Sons, Ltd, 1927

H. Poincaré, Méthodes Nouvelles de la Mécanique Céleste, Paris 1892-99

COURS DE L'ÉCOLE D'ÉTÉ DE PHYSIQUE THÉORIQUE

QUANTUM MECHANICS

Bryce S. DEWITT



★ ★

LES HOUCHES
ÉTÉ 1953

Q U A N T U M M E C H A N I C S

Lecture Notes by

Bryce Seligman DeWITT
Department of Physics
University of California
Berkeley

Chapter III

~~Appendix~~

Cours professé à
l'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie), France
Juillet 1953

3. THE INTERACTION OF MATTER AND RADIATION. CLASSICAL THEORY

Historically, the formalism of the preceding chapters was developed to deal with the problem of the motions of material particles (or bodies composed of material particles) under the direct influence of mutual (or external) "actions at a distance". The most well known law of force is that of Newtonian gravitation, which is expressed by the equations of motion

$$m_i \ddot{\mathbf{r}}_i = -G \sum_{j \neq i} m_i m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}, \quad N S, \tag{3.1}$$

where $\mathbf{r}_i$ is the position and m_i the mass of the i th gravitating particle, and G is the universal gravitational constant. The equations (3.1) can be derived from a Lagrangian function of the form

$$L = \sum_i \frac{1}{2} m_i \dot{r}_i^2 + \frac{1}{2} G \sum_{i,j}{}' \frac{m_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} . \tag{3.2}$$

Despite their apparent simplicity, the historical importance of Eqs. (3.1) should not be underestimated. It would be safe to say that the search for general solutions of these equations has stimulated the development of a very appreciable amount of modern mathematics, to say nothing of physics itself. In fact, the totality of consequences of these equations is still not known. (For a full, up-to-date account of nearly all that is known about these equations, see Aurel Wintner, The Analytical Foundations of Celestical Mechanics, Princeton University Press, 1941.)

However, in a sense it may be said that Eqs. (3.1) are well understood and their consequences fairly completely known. For, there is thought to be no conceptual difficulty lying hidden

in these equations nor any essential analytical difficulty remaining in their application to special problems. In fact, they have been so completely successful in correlating the phenomena of celestical mechanics that they can easily be solved to an accuracy which has no validity because it enters into the domain governed by relativistic effects.

Far less is all this true of the dynamical equations describing the motions of material particles which interact, not directly, but through the intermediary of "fields". Although the methods of the preceding chapters are also applicable to field-particle systems, the conceptual and analytical difficulties presented by theses systems are so formidable that in only two cases, namely that of electrodynamics and that of Einsteinian gravitation, have the systems been of such an elegant character and the physical principles involved so clearly defined that it has been found possible to isolate the difficulties in such a manner as at least partially to resolve them.. The resulting theories in these two cases are actually in amazingly good agreement with experiment, considering how poorly they are understood compared to the Newtonian theory of gravitation. And this is a very fortunate circumstance, because a well developed classical theory of radiation was absolutely essential to the development of the modern quantum theory.

We therefore turn next to a study of the mathematical essentials of the classical theory of the interaction of matter and radiation. We shall use a somewhat more elegant method than was actually available at the time of the birth of the quantum theory (or even of quantum mechanics itself). It has the advantage of compactness, modernity, and manifest Lorentz invariance throughout. The method is due to Dirac (P.A.M. Dirac, Proc. Roy. Soc., A.167 (1938), p. 148.). Following well established precedent we shall ignore the Einsteinian gravitational field, as gravitation has never been observed to take part in physical events on a quantum level, and we here have our eye only on later applications to quantum theory.

Consider the system composed of a single material particle of mass m, carrying an electric charge e , and interacting with an

electromagnetic field $F_{\mu\nu}$. Its motion, and that of the field it produces, are experimentally observed to be described by a Lagrangian density function of the form

$$\mathcal{L}(x) = -mc^2 \int_{-\infty}^{+\infty} (-\dot{z}^2)^{\frac{1}{2}} \delta(x-z)\, d\tau + e \int_{-\infty}^{+\infty} A_\mu(x)\, \dot{z}_\mu\, \delta(x-z)\, d\tau$$
$$- \frac{1}{16\pi} F_{\mu\nu}(x)\, F_{\mu\nu}(x)$$
$$= c \int_{-\infty}^{+\infty} L_0\, \delta(x-z)\, d\tau + \frac{1}{c} j_\mu(x)\, A_\mu(x) - \frac{1}{16\pi} F_{\mu\nu}(x)\, F_{\mu\nu}(x) \qquad (3.3.)$$

where

$$L_0 \equiv -mc(-\dot{z}^2)^{\frac{1}{2}} \qquad (3.4.)$$

and

$$j_\mu(x) \equiv ec \int_{-\infty}^{+\infty} \dot{z}_\mu\, \delta(x-z)\, d\tau \qquad (3.5.)$$

Here z_μ is the space-time position of the particle, its space-time path being given in parametric form by functions $z_\mu(\tau)$ of a parameter τ . For the time being τ is completely arbitrary (except that it must increase with z_0) but it will later be convenient to take take it as the proper time of the particle. The dot denotes differentiation with request to τ . C is the velocity of light, A_μ is the electromagnetic vector potential, j_μ is the current 4-vector, and L_0 is the Lagrangian function of the free particle. We have

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu} \quad , \qquad (3.6.)$$

where commas followed by indices denote differentiation with respect to the space-time coordinates x_μ . In non-invariant notation we have

$$(A_\mu) = (A, i\varphi) \qquad (3.7.)$$

$$F_{\mu\nu} = \begin{pmatrix} 0 & H_3 & -H_2 & -iE_1 \\ -H_3 & 0 & H_1 & -iE_2 \\ H_2 & -H_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{pmatrix} \tag{3.8}$$

$$E = -\nabla\varphi - \frac{1}{c}\frac{\partial A}{\partial t} \tag{3.9}$$

$$H = \nabla \times A \tag{3.10}$$

$$(j_\mu) = (j, ic\rho) \tag{3.11}$$

where A is the 3-vector potential, φ is the "scalar" potential, E is the electric field vector, H is the magnetic field vector, j is the current 3-vector, and ρ is the charge density.

current conservation

The current 4-vector, when expressed in the form (3.5), is readily seen to satisfy automatically the charge conservation equation :

$$j_{\mu,\mu}(x) = ec\int_{-\infty}^{+\infty} \dot{z}_\mu \, \delta_{,\mu}(x-z)\, d\tau$$

$$= -ec\int_{-\infty}^{+\infty} \frac{d}{d\tau}\,\delta(x-z)\, d\tau = -ec\Big[\delta(x-z)\Big]_{\tau=-\infty}^{+\infty} = 0 \tag{3.12}$$

Problem XXIII : Show that in non-invariant notation (3.5) reduces to

$$j(r,t) = e\, v(t)\, \delta[r - \zeta(t)]$$

and $$\rho(r,t) = e\,\delta[r - \zeta(t)]$$

where $\zeta(t)$ is the 3-vector position of the particle as a function of the time and $v(t)$ is the velocity of the particle :

$$v(t) = \zeta'(t)$$

the prime here denoting differentiation with respect to the time.

---

equation of motion

The equations of motion are obtained from the variational principle

$$0 = \frac{1}{c}\,\delta \int \mathcal{L}(x)\,dx$$

$$0 = \int_{-\infty}^{+\infty} \delta L_0\,d\tau + \frac{e}{c}\int_{\infty}^{\infty}\int_{-\infty}^{+\infty} A_\mu(x)\left[\delta \dot z_\mu\,\delta(x-z) - \dot z_\mu\,\delta_{,\nu}(x-z)\,\delta z_\nu\right] d\tau\,dx$$

$$+ \frac{1}{c^2}\int_\infty f_\mu\,\delta A_\mu\,dx + \frac{1}{4\pi c}\int_\infty F_{\mu\nu}\,\delta A_{\mu,\nu}\,dx$$

$$= \int_{-\infty}^{\infty}\Big\{ mc\left(-\dot z^2\right)^{\frac12}\dot z_\mu\,\delta z_\mu + \frac{e}{c}\int_\infty \Big[A_\mu(x)\,\delta z_\mu\,\delta_{,\nu}(x-z)\,\dot z_\nu - A_\nu(x)\,\dot z_\nu\,\delta_{,\mu}(x-z)\,\delta z_\mu\Big]dx\Big\}\,d\tau$$

$$+ \frac{1}{c}\int_\infty \left[-\frac{1}{4\pi}F_{\mu\nu,\nu} + \frac{1}{c} f_\mu\right]\delta A_\mu\,dx$$

$$= \int_{-\infty}^{+\infty}\left\{-mc\,\frac{d}{d\tau}\left[\left(-\dot z^2\right)^{\frac12}\dot z_\mu\right] + \frac{e}{c}\,F_{\mu\nu}(z)\,\dot z_\nu\right\}\delta z_\mu\,d\tau$$

$$+ \frac{1}{c}\int_\infty\left\{-\frac{1}{4\pi}\,F_{\mu\nu,\nu} + \frac{1}{c}\,f_\mu\right\}\delta A_\mu\,dx \tag{3.13}$$

Choosing τ to be the proper time, so that

$$\dot z^2 = -c^2\,, \tag{3.14}$$

we have the familiar equations

$$m\ddot z_\mu = \frac{e}{c}\,F_{\mu\nu}(z)\,\dot z_\nu\,, \tag{3.15}$$

$$F_{\mu\nu,\nu} = \frac{4\pi}{c}\,f_\mu\,. \tag{3.16}$$

Although equation (3.15) is the correct equation to give the motion of a charged particle in an impressed electromagnetic field, and equation (3.16) is the correct equation to give the production of an electromagnetic field by an impressed current, together they lead to a mathematical dilemma. As is well known, infinities lie hidden in the equations (3.15, 3.16). In particular, the quantity $F_{\mu\nu}(z)$ appearing on the right of (3.15) is divergent when given by (3.16). Therefore we shall presently ignore Eq. (3.15) and concentrate on Eq. (3.16).

Stress-Tensor

First, however, we shall pretend innocence and treat both equations as entirely respectable in order to investigate the distribution of energy and momentum in the system. Now, the momentum 3-vector of the particle is given by

$$p_i^{\mathbf{p}} = m \dot{z}_i \,, \tag{3.17}$$

differentiation being with respect to the proper time. The force acting on the particle is therefore

$$F_i = \frac{d}{dt} p_i^{\mathbf{p}} = \frac{d\tau}{dt} m \ddot{z}_i \,, \tag{3.18}$$

and the rate at which energy is imparted to the particle is given by

$$\frac{dE^{\mathbf{p}}}{d\tau} = F_i \dot{z}_i = \frac{d\tau}{dt} m \ddot{z}_i \dot{z}_i = \frac{d\tau}{dt} m \ddot{z}_0 \dot{z}_0 = mc \ddot{z}_0 \tag{3.19}$$

where we have used the identity

$$\dot{z}_\mu \dot{z}_\mu = 0 \,, \tag{3.20}$$

which can be obtained by differentiating Eq. (3.14), and the fact that $z_0 = ct$. The energy of the particle can evidently be identified with $mc \dot{z}_0$. Introducing the energy momentum 4-vector

$$p_\mu^{\mathbf{p}} = m \dot{z}_\mu \tag{3.21}$$

we may write

$$(p_\mu^{\mathbf{p}}) = \left(p^{\mathbf{p}}, \frac{i}{c} E^{\mathbf{p}}\right) \tag{3.22}$$

For a free particle we may write

$$P_\mu^P = \frac{1}{c} \int_\sigma T_{\mu\nu}^P \, d\sigma_\nu \tag{3.23}$$

where σ is an arbitrary space-like surface and where

$$T_{\mu\nu}^P(x) = mc \int_{-\infty}^{+\infty} \dot{z}_\mu \dot{z}_\nu \, \delta(x-z) \, d\tau \, . \tag{3.24}$$

For, on substituting (3.24) into (3.23) we have

$$P_\mu^P = m \int_\sigma \int_{-\infty}^{+\infty} \dot{z}_\mu \dot{z}_\nu \, \delta(x-z) \, d\tau \, d\sigma_\nu \, . \tag{3.25}$$

Since $\dot{z}_\mu$ is constant it can be moved outside the integral signs, and then the obvious relation

$$\int_\sigma \int_{-\infty}^{+\infty} \dot{z}_\nu \, \delta(x-z) \, d\tau \, d\sigma_\nu = 1 \tag{3.26}$$

can be used to obtain (3.21). Since, for a free particle, P_μ^P is constant, expression (3.23) is independent of the space-like surface σ, which implies $T_{\mu\nu,\nu}^P = 0$ (See Eq. (3.11) of the Appendix.) In the present case, however, we have, using (3.15) and (3.16)

$$T_{\mu\nu,\nu}^P = mc \int_{-\infty}^{\infty} \dot{z}_\mu \dot{z}_\nu \, \delta_{,\nu}(x-z) \, d\tau = -mc \int_{-\infty}^{+\infty} \dot{z}_\mu \frac{d}{d\tau} \delta(x-z) \, d\tau$$

$$= mc \int_{-\infty}^{+\infty} \ddot{z}_\mu \, \delta(x-z) \, d\tau \qquad = e \int_{-\infty}^{+\infty} F_{\mu\nu}(z) \, \dot{z}_\nu \, \delta(x-z) \, d\tau$$

$$= \frac{1}{c} F_{\mu\nu}(x) J_\nu(x) = \frac{1}{4\pi} F_{\mu\nu}(x) F_{\nu\sigma,\sigma}(x) \tag{3.27}$$

But

$$F_{\mu\nu} F_{\nu\sigma,\sigma} = -F_{\mu\sigma} F_{\nu\sigma,\nu} = -\left(F_{\mu\sigma} F_{\nu\sigma}\right)_{,\nu} + F_{\mu\sigma,\nu} F_{\nu\sigma}$$

and

$$F_{\mu\sigma,\nu} F_{\nu\sigma} = \frac{1}{2}\left(F_{\mu\sigma,\nu} + F_{\nu\mu,\sigma}\right) F_{\nu\sigma} = \frac{1}{2} F_{\nu\sigma,\mu} F_{\nu\sigma} = \frac{1}{4}\left(F_{\nu\sigma} F_{\nu\sigma}\right)_{,\mu}$$

Hence we may write

$$T_{\mu\nu,\nu}^P = -T_{\mu\nu,\nu}^F \tag{3.28}$$

where

$$T_{\mu\nu}^{F} = \frac{1}{4\pi}\left(F_{\mu\sigma} F_{\nu\sigma} - \frac{1}{4}\delta_{\mu\nu} F_{\sigma\tau} F_{\sigma\tau} \right) \tag{3.29}$$

$T_{\mu\nu}^{P}$ and $T_{\mu\nu}^{F}$ are known as the stress tensors of the particle and field respectively. If we introduce the total stress tensor

$$T_{\mu\nu} = T_{\mu\nu}^{P} + T_{\mu\nu}^{F} \tag{3.30}$$

we may express the conservation laws of energy and momentum in the compact form

$$T_{\mu\nu,\nu} = 0 \tag{3.31}$$

At points off the world line of the particle the tensor $T_{\mu\nu}^{P}$ vanishes and $T_{\mu\nu}$ reduces to the stress tensor $T_{\mu\nu}^{F}$ for the field. We shall be primarily interested in the value of $T_{\mu\nu}$ at points near to but not on the world line. Hence we shall turn our attention to $T_{\mu\nu}^{F}$, and to Eq. (3.16) which gives the law of generation of the field quantities $F_{\mu\nu}$ out of which $T_{\mu\nu}^{F}$ is constructed.

Solution of the equations.

If we choose a gauge in which the Lorentz condition

$$A_{\mu,\mu} = 0 \tag{3.32}$$

is satisfied, then Eq. (3.16) reduces to

$$\Box^2 A_\mu = -\frac{4\pi}{c} j_\mu \tag{3.33}$$

Equation (3.33) can readily be solved provided we know the Green's functions of the d'Alembertian operator $\Box^2$. Which Green's function we choose depends upon the boundary conditions which we set up for the problem. Our method here will be reminiscent of our treatment of the forced harmonic oscillator in chapter II. We shall replace equation (3.33) by the equation

$$\Box^2 A_\mu + 2\varepsilon_\nu A_{\mu,\nu} = -\frac{4\pi}{c} j_\mu \tag{3.34}$$

in which a dissipative term involving an infinitesiaml time-like vector ϵ_ν has been added. The vector ϵ_ν will be supposed to have a negative time-like orientation, so that it can be brought into the form $(\epsilon_\mu) = (0, 0, 0, -i\epsilon)$, with $\epsilon > 0$, by a certain proper Lorentz transformation. Since the 4-dimensional delta function is expressible in the Fourier form

$$\delta(x) = \frac{1}{(2\pi)^4} \int_\infty e^{ikx} d^4k \,, \tag{3.35}$$

where kx is short for $k_\mu x_\mu$, the Green's function of the operator $\Box^2 + 2\epsilon_\mu \frac{\partial}{\partial x_\mu}$ can immediately be written down as

$$G_{\Box^2 + 2\epsilon \frac{\partial}{\partial x}}(x) = -\frac{1}{(2\pi)^4} \int_\infty \frac{e^{ikx}}{k^2 - 2i\epsilon k} d^4k \tag{3.36}$$

Choosing a special coordinate system in which ϵ_μ lies along the time axis, we may write , using Eqs. (2.17, 18) of the Appendix,

$$\frac{1}{k^2 - 2i\epsilon_\mu k_\mu} = \frac{1}{k^2 - k_0^2 - 2i\epsilon k_0} = \frac{1}{2|k|}\left[\frac{1}{|k| + k_0 + i\epsilon} + \frac{1}{|k| - k_0 - i\epsilon}\right]$$

$$= -\frac{\pi i}{|k|}\left[\delta_+(|k| + k_0) - \delta_-(|k| - k_0)\right] \tag{3.37}$$

where

$$(k_\mu) = (k, ik_0) \tag{3.38}$$

Using Eqs: (2.15, 16) of the Appendix, we obtain, on substituting (3.37) into (3.36),

$$G_{\Box^2 + 2\epsilon \frac{\partial}{\partial x}}(x) = -\frac{\pi i}{(2\pi)^5} \int_\infty d^3k \int_0^\infty d\alpha \frac{1}{|k|}\left\{e^{i(|k| + k_0)\alpha} - e^{-i(|k| + k_0)\alpha}\right\} e^{ikx}$$

$$= \frac{\pi i}{(2\pi)^5} \int_\infty d^3k \int_{-\infty}^{\infty} dk_0 \int_0^\infty d\alpha \frac{1}{|k|} e^{ik_0(\alpha - x_0)}\left(e^{i|k|\alpha} - e^{-i|k|\alpha}\right) e^{ik\cdot z} \tag{3.39}$$

But

$$\int_\omega \frac{1}{|k|}\left(e^{+i|k|x} - e^{-i|k|x}\right) e^{ik\cdot r}\, d^3k$$

$$= 2\pi \int_0^\infty \int_{-1}^{+1} \frac{e^{ikrx}}{|k|}\left(e^{i|k|x} - e^{-i|k|x}\right) |k|^2\, dx\, d|k|$$

$$= -\frac{2\pi i}{r} \int_0^\infty \left(e^{i|k|x} - e^{-i|k|x}\right)\left(e^{i|k|r} - e^{-i|k|r}\right) d|k|$$

$$= -\frac{2\pi i}{r} \int_{-\infty}^{+\infty} \left[e^{i(x+r)|k|} - e^{i(x-r)|k|}\right] d|k|$$

$$= -\frac{4\pi^2 i}{r}\left[\delta(x+r) - \delta(x-r)\right] \tag{3.40}$$

Hence

$$G_{\nabla^2+k^2}\frac{\partial}{\partial x}(x) = \frac{1}{4\pi r}\int_0^\infty \delta(x-x_0)\left[\delta(x+r) - \delta(x-r)\right] dx$$

$$= -\frac{1}{4\pi r}\cdot\frac{1}{2}\left(1 + \frac{x_0}{|x_0|}\right)\delta(x_0 - r) \tag{3.41}$$

Equation (3.41) may be written in invariant form by observing that

$$\left(1 + \frac{x_0}{|x_0|}\right)\delta(x_0 + r) = 0. \tag{3.42}$$

since $r > 0$. Hence

$$G_{\nabla^2+k^2}\frac{\partial}{\partial x}(x) = -\frac{1}{4\pi r}\cdot\frac{1}{2}\left(1 + \frac{x_0}{|x_0|}\right)\left[\delta(x_0 + r) + \delta(x_0 - r)\right]$$

$$= -\frac{1}{4\pi}\left(1 + \frac{\epsilon x}{|\epsilon x|}\right)\delta(x^2), \tag{3.43}$$

in which we have made use of Eq. (1.25) of the Appendix. Since the

vector ϵ is completely arbitrary, save that it must be time-like, we may write more generally

$$G_{\Box^2 + \epsilon \frac{\partial}{\partial x}}(x) = -D^{ret}(x) \tag{3.44}$$

where

$$D^{ret}(x) = \frac{1}{4\pi}\left(1 + \varepsilon[x, \sigma_0]\right)\delta(x^2) \tag{3.45}$$

σ_0 is an arbitrary space-like surface passing through the origin, and the functimal $\varepsilon[x, \sigma_0]$ is equal to + 1 or - 1 according as x lies to the future or the past of σ_0 .

The function $D^{ret}(x)$ is known as the retarded Green's function of the d'Alembertian operator $\Box^2$. It vanishes at all points off the light cone through the origin, as well as at all points to the future of the origin. If we had inserted a negative dissipation term instead of a positive dissipation term in (3.34), we would have obtained the advanced Green's function of $\Box^2$:

$$D^{adv}(x) = \frac{1}{4\pi}\left(1 - \varepsilon[x, \sigma_0]\right)\delta(x^2) \tag{3.46}$$

The mean value of the retarded and advanced Green's functions is evidently the Green's function obtained in the Appendix, Eq. (1.60)

$$\bar{D}(x) = \frac{1}{2}\left[D^{ret}(x) + D^{adv}(x)\right] = \frac{1}{4\pi}\delta(x^2) \tag{3.47}$$

Another important invariant function is the difference of the advanced and retarded Green's functions :

$$D(x) \equiv D^{adv}(x) - D^{ret}(x) = -\frac{1}{2\pi}\varepsilon[x, \sigma_0]\delta(x^2) \tag{3.48}$$

It evidently satisfies the free wave equation

$$\Box^2 D(x) = 0 \tag{3.49}$$

Returning now to Eq. (3.33), we see that a particular solution of it is given by

$$A_\mu^{ret}(x) = \frac{4\pi}{c} \int_\infty D^{ret}(x-x')\, j_\mu(x')\, dx' \tag{3.50}$$

A_μ^{ret} is called the retarded potential of the current j_μ. The general solution of (3.33) is

$$A_\mu = A_\mu^{ret} + A_\mu^{in} \tag{3.51}$$

where A_μ^{in} is an arbitrary "free" part satisfying the free wave equation

$$\Box^2 A_\mu^{in} = 0 \tag{3.52}$$

If we imagine the interaction between matter and radiation to be switched on adiabatically in the remote past, then A_μ^{in} may be interpreted as the potential of the incoming electromagnetic field which is subsequently to be scattered by the charged particle.

Inserting (3.5) and (3.45) into (3.50), we have

$$A_\mu^{ret}(x) = e \int_\infty dx' \int_{-\infty}^{+\infty} d\tau \left(1 + \varepsilon[x-x'; \sigma_0]\right) \delta\left((x-x')^2\right) \dot{z}_\mu\, \delta(x'-z) \tag{3.53}$$

But

$$\varepsilon[x-x'; \sigma_0] = -\varepsilon[x'; \sigma(x)] \tag{3.54}$$

where $\sigma(x)$ is the space-like surface obtained by displacing σ_0 in the time direction until it passes through x. Hence

$$\begin{aligned} A_\mu^{ret}(x) &= e \int_\infty dx' \int_{-\infty}^{+\infty} d\tau \left(1 - \varepsilon[x'; \sigma]\right) \delta\left((x-x')^2\right) \dot{z}_\mu\, \delta(x'-z) \\ &= 2e \int_{-\infty}^{\sigma(x)} dx' \int_{-\infty}^{+\infty} d\tau\, \dot{z}_\mu\, \delta\left((x-x')^2\right) \delta(x'-z) \\ &= 2e \int_{-\infty}^{\tau_6} \dot{z}_\mu\, \delta\left((x-z)^2\right) d\tau \end{aligned}$$

where τ_6 is the value of the parameter τ at the point of intersection of the particle world line with $\sigma(x)$. Introducing the abbreviations

$$\xi = (z-x)^2 \quad , \quad \xi_6 = \left(z(\tau_6) - x\right)^2 > 0 \quad , \tag{3.56}$$

So that

$$d\zeta = 2\dot{z}(z-x)\,d\tau\,, \tag{3.57}$$

we may write

$$A_\mu^{ret}(x) = e\int_{-\infty}^{\tau_c} \frac{\dot{z}_\mu}{\dot{z}(z-x)}\,\delta(\zeta)\,d\zeta$$

$$= e\,\frac{\dot{z}_\mu^2}{\dot{z}^2(z^2-x)} \tag{3.58}$$

where

$$(z^2-x)^2 = 0\,, \quad z_0^2 - x_0 < 0 \tag{3.59}$$

z_μ^2 is called the retarded position of the particle with respect to x.

The advanced potential is found in a similar manner to be given by

$$A_\mu^{adv}(x) = 2e\int_{\tau_c}^{\infty} \dot{z}_\mu\,\delta\left((x-z)^2\right)d\tau$$

$$= 2e\int_{\zeta_c}^{-\infty} \frac{\dot{z}_\mu}{\dot{z}(z-x)}\,\delta(\zeta)\,d\zeta$$

$$= -e\,\frac{\dot{z}_\mu^a}{\dot{z}^a(z^a-x)} \tag{3.60}$$

where

$$(z^a-x)^2 = 0\,, \quad z_0^a - x_0 > 0 \tag{3.61}$$

z_μ^a is called the advanced position of the particle with respect to x.

Problem XXIV : By writing (3.58) in . non-covariant form obtain the familiar Lienard-Wiechert potentials.

The retarded and advanced field strengths may also be readily found. First observe that, for $x \neq z^2$,

$$A^{ret}_{\mu,\nu}(x) = 4e \int_{-\infty}^{\tau_6} \dot{z}_\mu (x_\nu - z_\nu)\, \delta'((x-z)^2)\, d\tau$$

$$= 2e \int_{-\infty}^{\tau_6} \frac{\dot{z}_\mu (x_\nu - z_\nu)}{\dot{z}(z-x)} \frac{d}{d\tau} \delta((x-z)^2)\, d\tau$$

$$= -2e \int_{-\infty}^{\tau_6} \left[\frac{d}{d\tau} \frac{\dot{z}_\mu (x_\nu - z_\nu)}{\dot{z}(z-x)}\right] \delta((x-z)^2)\, d\tau$$

$$= -e \int_{-\infty}^{s_6} \frac{1}{\dot{z}(z-x)} \left[\frac{d}{d\tau} \frac{\dot{z}_\mu (x_\nu - z_\nu)}{\dot{z}(z-x)}\right] \delta(s)\, ds \qquad (3.62)$$

Hence

$$F^{ret}_{\mu\nu}(x) = A^{ret}_{\nu,\mu}(x) - A^{ret}_{\mu,\nu}(x)$$

$$= -e \frac{\ddot{z}^2_\mu (z^2_\nu - x_\nu) - \ddot{z}^2_\nu (z^2_\mu - x_\mu)}{\left[\dot{z}^2 (z^2 - x)\right]^2}$$

$$+ e \frac{\left[\dot{z}^2_\mu (z^2_\nu - x_\nu) - \dot{z}^2_\nu (z^2_\mu - x_\mu)\right]\left[\ddot{z}^2 (z^2 - x) + (\dot{z}^2)^2\right]}{\left[\dot{z}^2 (z^2 - x)\right]^3} \qquad (3.63)$$

Similarly

$$F^{adv}_{\mu\nu}(x) = A^{adv}_{\nu,\mu}(x) - A^{adv}_{\mu,\nu}(x)$$

$$= -e\,\frac{\ddot{z}^a_\mu\left(z^a_\nu - x_\nu\right) - \ddot{z}^a_\nu\left(z^a_\mu - x_\mu\right)}{\left[\dot{z}^a\left(z^a - x\right)\right]^2}$$

$$+ e\,\frac{\left[\dot{z}^a_\mu\left(z^a_\nu - x_\nu\right) - \dot{z}^a_\nu\left(z^a_\mu - x_\mu\right)\right]\left[\ddot{z}^a\left(z^a - x\right) + \left(\dot{z}^a\right)^2\right]}{\left[\dot{z}^a\left(z^a - x\right)\right]^3} \qquad (3.64)$$

In analogy with Eq. (3.51) we may write

$$A_\mu = A^{adv}_\mu + A^{out}_\mu \qquad (3.65)$$

where A^{out}_μ, like A^{in}_μ, satisfies the free wave equation

$$\Box^2 A^{out}_\mu = 0 \qquad (3.66)$$

A^{out}_μ may be interpreted as the potential of the outgoing electromagnetic field after a scattering process has taken place. If we define the mean generated potential by

$$\bar{A}_\mu(x) = \frac{4\pi}{c}\int_\infty \bar{D}(x-x')\,j_\mu(x')\,dx' = \frac{1}{2}\left[A^{ret}_\mu(x) + A^{adv}_\mu(x)\right] \qquad (3.67)$$

then we may write

$$A_\mu = \bar{A}_\mu + \bar{A}^{free}_\mu \qquad (3.68)$$

where

$$\bar{A}^{free}_\mu = \frac{1}{2}\left(A^{in}_\mu + A^{out}_\mu\right), \quad \Box^2 \bar{A}^{free}_\mu = 0 \qquad (3.69)$$

It will also be convenient to write

$$\bar{A}_\mu^{\text{free}} = A_\mu^{\text{in}} + \tfrac{1}{2}\left(A_\mu^{\text{out}} - A_\mu^{\text{in}}\right) = A_\mu^{\text{in}} + \tfrac{1}{2}\left(A_\mu^{\text{ret}} - A_\mu^{\text{adv}}\right)$$

$$= A_\mu^{\text{in}} + \tfrac{1}{2} A_\mu^{\text{rad}} \tag{3.70}$$

where

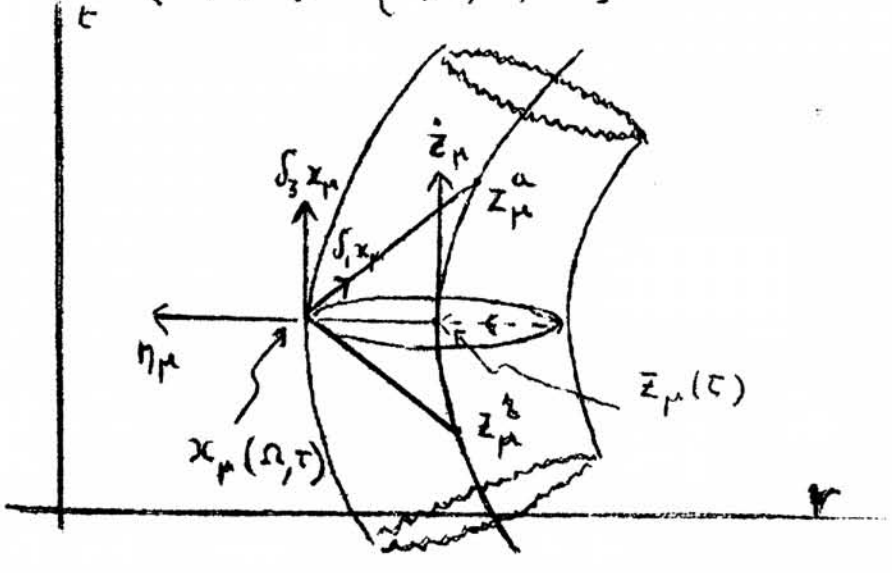
$$A_\mu^{\text{rad}}(x) = A_\mu^{\text{ret}}(x) - A_\mu^{\text{adv}}(x) = -\frac{4\pi}{c}\int_\infty D(x-x') j_\mu(x')\, dx' \tag{3.71}$$

A_μ^{rad} will be seen later to have a special significance in electrodynamic theory.

We shall be interested in evaluating the electrodynamic stress tensor, and hence expressions (3.63, 64), at points near the world line of the particle. More precisely, we shall be interested in evaluating the stress tensor at points on a tube in space-time which encloses the particle world line. This tube will be constructed in the following manner : (See diagram).

Consider a particular point $z_\mu(\tau)$ on the world line of the particle. Let us carry out a Lorentz transformation to a coordinate system in which the particle is momentarily at rest at that point. Then, if $\tilde{\tau}$ is taken to be the proper time, we shall have

$$\left(\dot{\xi}_\mu(\tau)\right) = (0, 0, 0, ic) \tag{3.72}$$



Construct a sphere of radius ϵ centered at the point $z_\mu(\tau)$ and lying in the space-like cross section (t = constant) through $z_\mu(t)$ The points on this sphere may be labeled by the value of τ in question and by the element of spherical solid angle Ω to which they happen to correspond. Explicitly we may write $x_\mu(\Omega, \tau)$, although frequently we shall omit the labels Ω and τ and write simply x_μ the relation between x_μ and z_μ being understood. The locus of all points $x_\mu(\Omega,\tau)$, for all values of the parameters Ω , τ , constitutes the tube in space-time surrounding the particle world line. The spherical radius ϵ is constant and independent of τ , and we shall subsequently be interested in the limit in which ϵ becomes infinitesimally small.

Now, from construction, we may write

$$(x - z)^2 = \epsilon^2 , \tag{3.73}$$

$$\dot{z}(x - z) = 0 \tag{3.74}$$

Let δx_μ be an arbitrary displacement of the point x_μ tangent to the tube, and let $\delta\tau$ be the corresponding variation in the proper time. Then

$$0 = (x_\mu - z_\mu)(\delta x_\mu - \dot{z}_\mu d\tau) = (x_\mu - z_\mu)\delta x_\mu \tag{3.75}$$

$$0 = \ddot{z}(x - z)\delta\tau + \dot{z}_\mu(\delta x_\mu - \dot{z}_\mu \delta\tau)$$

$$= (c^2 + K)\delta\tau + \dot{z}_\mu \delta x_\mu \tag{3.76}$$

where

$$K = \ddot{z}(x - z) \tag{3.77}$$

K may be regarded as a measure of the curvature of the tube.

The displacement δx_μ may be regarded as a linear combination of three displacements $\delta_1 x_\mu$, $\delta_2 x_\mu$, $\delta_3 x_\mu$, where $\delta_1 x_\mu$ and $\delta_2 x_\mu$ correspond to a variation in Ω , while $\delta_3 x_\mu$ corresponds to a variation in τ (see diagram). In the rest system of the particle we have

$$(\delta_3 x_\mu) = (0, 0, 0, i\delta_3 x_0) \tag{3.78}$$

where

$$\delta_3 x_0 = c^{-1}\dot{z}_0\,\delta x_0 = -c^{-1}\dot{z}_\mu\,\delta x_\mu = c^{-1}(c^2+\kappa)\,\delta\tau \tag{3.79}$$

Hence

$$-\delta\ell^2 = \delta_3 x_\mu\,\delta_3 x_\mu = -c^{-2}(c^2+\kappa)^2\,\delta\tau^2 \tag{3.80}$$

where $\delta\ell$ is the element of lenght along the tube, corresponding to the variation $\delta\tau$.

The three displacments $\delta_1 x_\mu$, $\delta_2 x_\mu$, $\delta_3 x_\mu$ may be regarded as defining an element of surface $d\sigma$ on the tube, given by

$$d\sigma = \epsilon^2\, d\Omega\, d\ell = c^{-1}\epsilon^2(c^2+\kappa)\, d\Omega\, d\tau \tag{3.81}$$

It will also be convenient to introduce the directed surface element defined by

$$d\sigma_\mu = \eta_\mu\, d\sigma = c^{-1}\epsilon(c^2+\kappa)(x_\mu - z_\mu)\, d\Omega\, d\tau \tag{3.82}$$

where η_μ is the unit outward normal vector to the tube :

$$\eta_\mu = \epsilon^{-1}(x_\mu - z_\mu) \tag{3.83}$$

Now, in order to evaluate expressions (3.63, 64) for $F^{ret}_{\mu\nu}$ and $F^{adv}_{\mu\nu}$ at the point x_μ we must express the points z^r_μ and z^a_μ in terms of z_μ and x_μ . To do this we must have recourse to the method of series expansion. Denoting the retarded and advanced proper times respectively by τ_r and τ_a , let

us introduce the quantities

$$\sigma_+ = \tau_a - \tau > 0 \quad , \quad \sigma_- = \tau_2 - \tau < 0 \tag{3.84}$$

Generally, in what follows + signs will refer to advanced quantities and - signs to retarded quantities. We have

$$z^{\pm} = z_\mu + \sigma_\pm \dot{z}_\mu + \frac{1}{2}\sigma_\pm^2 \ddot{z}_\mu + \frac{1}{6}\sigma_\pm^3 \dddot{z}_\mu + \cdots \tag{3.85}$$

Moreover

$$0 = (x - z^{\pm})^2$$

$$= \left(\xi - \sigma_\pm \dot{z} - \frac{1}{2}\sigma_\pm^2 \ddot{z} - \frac{1}{6}\sigma_\pm^3 \dddot{z} - \cdots \right)^2 \tag{3.86}$$

where

$$\xi_\mu \equiv x_\mu - z_\mu$$

Remembering the identities

$$\left.\begin{aligned} \dot{z}_\mu \dot{z}_\mu &= -c^2 \\ \dot{z}_\mu \ddot{z}_\mu &= 0 \\ \dot{z}_\mu \dddot{z}_\mu &= -\ddot{z}^2 \\ \dot{z}_\mu \xi_\mu &= 0 \\ \ddot{z}_\mu \xi_\mu &= K \\ \therefore \xi_\mu \xi_\mu &= \epsilon^2 \end{aligned}\right\} \tag{3.87}$$

we may write (3.86) in the form

$$0 = \epsilon^2 - c^2\sigma_\pm^2 + \frac{1}{4}\sigma_\pm^4 \ddot{z}^2 + \cdots$$
$$- K\sigma_\pm^2 - \frac{1}{3}\sigma_\pm^3 \xi_\mu \dddot{z}_\mu - \frac{1}{3}\sigma_\pm^4 \ddot{z}^2$$

$$= \epsilon^2 - (c^2 + K)\sigma_\pm^2 - \frac{1}{12}\sigma_\pm^4 \ddot{z}^2 - \frac{1}{3}\sigma_\pm^3 \xi_\mu \dddot{z}_\mu + \cdots \tag{3.88}$$

correct to the 4th order in small quantities. Eq. (3.88) may be readily solved for $\sigma_{\pm}^2$ correct to the 4th order :

$$\sigma_{\pm}^2 = S^{-2}\left(\epsilon^2 - \frac{1}{12} S^{-4}\epsilon^4 \ddot{z}^{\cdot 2} \mp \frac{1}{3} S^{-3}\epsilon^3 \xi_\mu \dddot{z}_\mu + \cdots\right) \tag{3.89}$$

where we have introduced the abbreviation

$$S \equiv \left(c^2 + \kappa\right)^{\frac{1}{2}} \tag{3.90}$$

Extracting the square root of Eq. (3.89) and taking note of (3.84), we obtain

$$\sigma_{\pm} = \pm S^{-1}\epsilon\left(1 - \frac{1}{24} S^{-4}\epsilon^2 \ddot{z}^{\cdot 2} \mp \frac{1}{6} S^{-3}\epsilon\, \xi_\mu \dddot{z}_\mu + \cdots\right) \tag{3.91}$$

correct to the 3rd order.

Now we are able to write

$$\begin{aligned} z_\mu^{\pm} - x_\mu &= -\xi_\mu + \sigma_{\pm}\dot{z}_\mu + \frac{1}{2}\sigma_{\pm}^2 \ddot{z}_\mu + \frac{1}{6}\sigma_{\pm}^3 \dddot{z}_\mu + \cdots \\ &= -\xi_\mu \pm S^{-1}\epsilon\,\dot{z}_\mu\left(1 - \frac{1}{24} S^{-4}\epsilon^2 \ddot{z}^{\cdot 2} \mp \frac{1}{6} S^{-3}\epsilon\, \xi_\nu \dddot{z}_\nu\right) \\ &\quad + \frac{1}{2} S^{-2}\epsilon^2 \ddot{z}_\mu \pm \frac{1}{6} S^{-3}\epsilon^3 \dddot{z}_\mu + \cdots \end{aligned} \tag{3.92}$$

to the 3rd order

$$\begin{aligned} \dot{z}_\mu^{\pm} &= \dot{z}_\mu + \sigma_{\pm}\ddot{z}_\mu + \frac{1}{2}\sigma_{\pm}^2 \dddot{z}_\mu + \cdots \\ &= \dot{z}_\mu \pm S^{-1}\epsilon\,\ddot{z}_\mu + \frac{1}{2} S^{-2}\epsilon^2 \dddot{z}_\mu + \cdots \end{aligned} \tag{3.92}$$

to the 2nd order, and

$$\ddot{z}_\mu^{\pm} = \ddot{z}_\mu + \sigma_{\pm}\dddot{z}_\mu + \cdots = \ddot{z}_\mu \pm S^{-1}\epsilon\,\dddot{z}_\mu + \cdots \tag{3.93}$$

to the 1st order. Thus

$$\dot{z}^{\pm}(z-x) = \mp s^{-1}\epsilon c^2\left(1-\frac{1}{24}s^{-4}\epsilon^2\ddot{z}^2 \mp \frac{1}{6}s^{-3}\epsilon\xi_\mu\dddot{z}_\mu\right)\mp\frac{1}{6}s^{-3}\epsilon^3\ddot{z}^2$$
$$\mp s\epsilon K \pm \frac{1}{2}s^{-3}\epsilon^3\ddot{z}^2 - \frac{1}{2}s^{-2}\epsilon^2\xi_\mu\ddot{z}_\mu \mp \frac{1}{2}s^{-3}\epsilon^3\ddot{z}^2+\cdots$$

$$= \mp s\epsilon \mp \frac{1}{8}s^{-3}\epsilon^3\ddot{z}^2 - \frac{1}{3}s^{-2}\epsilon^2\xi_\mu\dddot{z}_\mu+\cdots \qquad (3.94)$$

to 3rd order,

$$\left[\dot{z}^{\pm}(z-x)\right]^{-2} = s^{-2}\epsilon^{-2}\left(1-\frac{1}{4}s^{-4}\epsilon^2\ddot{z}^2 \mp \frac{2}{3}s^{-3}\epsilon\xi_\mu\dddot{z}_\mu+\cdots\right) \qquad (3.95)$$

to - 2nd and 0th order,

$$\left[\dot{z}^{\pm}(z-x)\right]^{-3} = \mp s^{-3}\epsilon^{-3}\left(1-\frac{3}{8}s^{-4}\epsilon^2\ddot{z}^2 \mp s^{-3}\epsilon\xi_\mu\dddot{z}_\mu+\cdots\right) \qquad (3.96)$$

to - 3rd and - 1st order ,

$$\ddot{z}^{\pm}(z-x) = -K + \frac{1}{2}s^{-2}\epsilon^2\ddot{z}^2 \mp s^{-1}\epsilon\xi_\mu\dddot{z}_\mu - s^{-2}\epsilon^2\ddot{z}^2+\cdots \qquad (3.97)$$

$$\text{or}\quad \ddot{z}^{\pm}(z-x) - c^2 = -s^2 - \frac{1}{2}s^{-2}\epsilon^2\ddot{z}^2 \mp s^{-1}\epsilon\xi_\mu\dddot{z}_\mu+\cdots \qquad (3.98)$$

to 2nd order

$$\dot{z}^{\pm}\left(z^{\pm}_\nu - x_\nu\right) - \dot{z}_\nu^{\pm}\left(z^{\pm}_\mu - x_\mu\right)$$
$$= -\dot{z}_\mu\xi_\nu + \dot{z}_\nu\xi_\mu + \frac{1}{2}s^{-2}\epsilon^2\left(\dot{z}_\mu\ddot{z}_\nu - \dot{z}_\nu\ddot{z}_\mu\right) \pm \frac{1}{6}s^{-3}\epsilon^3\left(\dot{z}_\mu\dddot{z}_\nu - \dot{z}_\nu\dddot{z}_\mu\right)$$
$$\mp s^{-1}\epsilon\left(\ddot{z}_\mu\xi_\nu - \ddot{z}_\nu\xi_\mu\right) + s^{-2}\epsilon^2\left(\ddot{z}_\mu\dot{z}_\nu - \ddot{z}_\nu\dot{z}_\nu\right) - \frac{1}{2}s^{-2}\epsilon^2\left(\dddot{z}_\mu\xi_\nu - \dddot{z}_\nu\xi_\mu\right)$$
$$\pm\frac{1}{2}s^{-3}\epsilon^3\left(\dddot{z}_\mu\dot{z}_\nu - \dddot{z}_\nu\dot{z}_\mu\right)+\cdots$$
$$= -\left(\dot{z}_\mu\xi_\nu - \dot{z}_\nu\xi_\mu\right) - \frac{1}{2}s^{-2}\epsilon^2\left(\dot{z}_\mu\ddot{z}_\nu - \dot{z}_\nu\ddot{z}_\mu\right) \mp \frac{1}{3}s^{-3}\epsilon^3\left(\dot{z}_\mu\dddot{z}_\nu - \dot{z}_\nu\dddot{z}_\mu\right)$$
$$\mp s^{-1}\epsilon\left(\ddot{z}_\mu\xi_\nu - \ddot{z}_\nu\xi_\mu\right) - \frac{1}{2}s^{-2}\epsilon^2\left(\dddot{z}_\mu\xi_\nu - \dddot{z}_\nu\xi_\mu\right)+\cdots \qquad (3.99)$$

to 3 rd order,

$$\ddot{z}^{\pm}_{\mu}(z^{\mp}_{\nu} - x_{\nu}) - \ddot{z}^{\pm}_{\nu}(z^{\mp}_{\mu} - x_{\mu})$$
$$= -\ddot{z}_{\mu}\xi_{\nu} + \ddot{z}_{\nu}\xi_{\mu} \pm s^{-1}\epsilon(\ddot{z}_{\mu}\dot{z}_{\nu} - \ddot{z}_{\nu}\dot{z}_{\mu}) \mp s^{-1}\epsilon(\ddot{z}_{\mu}\xi_{\nu} - \ddot{z}_{\nu}\xi_{\mu})$$
$$+ s^{-2}\epsilon^{2}(\dddot{z}_{\mu}\dot{z}_{\nu} - \ddot{z}_{\nu}\dot{z}_{\mu}) + \cdots \qquad (3.100)$$

to 2nd order, and

$$\left[\ddot{z}^{\pm}(z - x) - c^{2}\right]\left[\dot{z}^{\mp}(z - x)\right]^{3}$$
$$= \mp s^{-3}\epsilon^{-3}\left(-s^{2} - \tfrac{1}{2}s^{-2}\epsilon^{2}\ddot{z}^{2} \mp s^{-1}\epsilon\,\xi_{\mu}\dddot{z}_{\mu} + \tfrac{3}{8}s^{-2}\epsilon^{2}\ddot{z}^{2} \pm s^{-1}\epsilon\,\xi_{\mu}\dddot{z}_{\mu} + \cdots\right)$$
$$= \mp s^{-3}\epsilon^{3}\left(-s^{2} - \tfrac{1}{8}s^{-2}\epsilon^{2}\ddot{z}^{2} + \cdots\right) \qquad (3.101)$$

to - 3rd and - 1st order.

We may now write, correct to - 2nd, - 1st; and 0th order,

$$F^{\pm}_{\mu\nu}(x)$$
$$= \pm e\,s^{-2}\epsilon^{-2}\Big\{-(\ddot{z}_{\mu}\xi_{\nu} - \ddot{z}_{\nu}\xi_{\mu}) \mp s^{-1}\epsilon(\dot{z}_{\mu}\ddot{z}_{\nu} - \dot{z}_{\nu}\ddot{z}_{\mu})$$
$$\mp s^{-1}\epsilon(\dddot{z}_{\mu}\xi_{\nu} - \ddot{z}_{\nu}\xi_{\mu}) - s^{-2}\epsilon^{2}(\dot{z}_{\mu}\dddot{z}_{\nu} - \dot{z}_{\nu}\ddot{z}_{\mu})\Big\}$$
$$+ e\,s^{-3}\epsilon^{-3}\Big\{s^{2}(\dot{z}_{\mu}\xi_{\nu} - \dot{z}_{\nu}\xi_{\mu}) + \tfrac{1}{8}s^{-2}\epsilon^{2}\ddot{z}^{2}(\dot{z}_{\mu}\xi_{\nu} - \dot{z}_{\nu}\xi_{\mu})$$
$$+ \tfrac{1}{2}\epsilon^{2}(\dot{z}_{\mu}\ddot{z}_{\nu} - \dot{z}_{\nu}\ddot{z}_{\mu}) \pm \tfrac{1}{3}s^{-1}\epsilon^{3}(\dot{z}_{\mu}\dddot{z}_{\nu} - \dot{z}_{\nu}\dddot{z}_{\mu})$$
$$\pm s\epsilon(\ddot{z}_{\mu}\xi_{\nu} - \ddot{z}_{\nu}\xi_{\mu}) + \tfrac{1}{2}\epsilon^{2}(\dddot{z}_{\mu}\xi_{\nu} - \ddot{z}_{\nu}\xi_{\mu})\Big\}$$

(3.102)

From Eq. (3.102) we may obtain

$$\overline{F_{\mu\nu}}(x) = \frac{1}{2}\left[F_{\mu\nu}^{ret}(x) + F_{\mu\nu}^{adv}(x)\right]$$

$$= e\left\{-\frac{1}{2}S^{-3}\epsilon^{-1}\left(\dot{z}_\mu \ddot{z}_\nu - \dot{z}_\nu \ddot{z}_\mu\right) - \frac{1}{2}S^{-3}\epsilon^{-1}\left(\dddot{z}_\mu \xi_\nu - \ddot{z}_\nu \xi_\mu\right)\right.$$

$$\left. + \left(S^{-1}\epsilon^{-3} + \frac{1}{8}S^{-5}\epsilon^{-1}\ddot{z}^2\right)\left(\dot{z}_\mu \xi_\nu - \dot{z}_\nu \xi_\mu\right)\right\} \qquad (3.103)$$

go to (3.124)

correct to the - 2nd, - 1st, and 0th orders in small quantities.

The mean generated field $\overline{F_{\mu\nu}}$ is seen to become infinite on the world line of the particle, i.e. as $\epsilon \to 0$. The field $F_{\mu\nu}^{real}$ defined by (3.71) ; on the other hand, turns out to be finite everywhere. Remembering that $S \to c$ as $\epsilon \to 0$, we find immediatly from (3.102)

$$F_{\mu\nu}^{real}(z) = \lim_{\epsilon \to 0}\left[\overline{F}_{\mu\nu}^{ret}(x) - F_{\mu\nu}^{adv}(x)\right]$$

$$= \frac{4}{3} e c^{-4}\left(\dot{z}_\mu \dddot{z}_\nu - \dot{z}_\nu \dddot{z}_\mu\right) . \qquad (3.104)$$

Let us now consider the integral

$$\frac{1}{c}\int_{\tau_1}^{\tau_2} T_{\mu\nu}\, d\sigma_\nu \qquad (3.105)$$

taken over all points of the tube lying between the proper time values τ_1 and τ_2 . On the tube we have

$$T_{\mu\nu}\, d\sigma_\nu = T_{\mu\nu}^{F} d\sigma_\nu = \frac{1}{4\pi}\left(F_{\mu\sigma} F_{\nu\sigma}\, d\sigma_\nu - \frac{1}{4} F_{\sigma\tau} F_{\sigma\tau}\, d\sigma_\mu\right)$$

$$= \frac{1}{4\pi}\left(\overline{F}_{\mu\sigma}\overline{F}_{\nu\sigma}\, d\sigma_\nu + \overline{F}_{\mu\sigma}^{free}\overline{F}_{\nu\sigma}\, d\sigma_\nu + \overline{F}_{\mu\sigma}\overline{F}_{\nu\sigma}^{free} d\sigma_\nu\right.$$

$$\left. - \frac{1}{4}\overline{F}_{\sigma\tau}\overline{F}_{\sigma\tau}\, d\sigma_\mu - \frac{1}{2}\overline{F}_{\sigma\tau}^{free}\overline{F}_{\sigma\tau}\, d\sigma_\mu + \cdots\right) \qquad (3.106)$$

where we have retained only those terms of the -2nd, - 1st, and 0th order in small quantities. But, using the identities

$$(\dot{z}_\mu \ddot{z}_6 - \dot{z}_6 \ddot{z}_\mu)(\dot{z}_\nu \ddot{z}_6 - \dot{z}_6 \ddot{z}_\nu) = \ddot{z}^2 \dot{z}_\mu \dot{z}_\nu - \cdot c^2 \ddot{z}_\mu \ddot{z}_\nu , \qquad (3.107)$$

$$(\dot{z}_\mu \xi_6 - \dot{z}_6 \xi_\mu)(\dot{z}_\nu \xi_6 - \dot{z}_6 \xi_\nu) = \epsilon^2 \dot{z}_\mu \dot{z}_\nu - c^2 \xi_\mu \xi_\nu \qquad (3.108)$$

$$(\dot{z}_\mu \ddot{z}_6 - \dot{z}_6 \ddot{z}_\nu)(\dot{z}_\nu \xi_6 - \dot{z}_6 \xi_\nu) = K \dot{z}_\mu \dot{z}_\nu - c^2 \ddot{z}_\mu \xi_\nu \qquad (3.109)$$

$$(\dddot{z}_\mu \xi_6 - \ddot{z}_6 \xi_\mu)(\dot{z}_\nu \xi_6 - \dot{z}_6 \xi_\nu) = \epsilon^2 \dddot{z}_\mu \dot{z}_\nu - \xi_6 \ddot{z}_6 \xi_\mu \dot{z}_\nu - \dot{z}^2 \xi_\mu \xi_\nu \qquad (3.110)$$

and remembering

$$d\sigma_\mu = c^{-1} s^2 \epsilon \xi_\mu \, d\Omega \, d\tau \qquad (3.111)$$

we may write

$$\begin{aligned} \overline{F}_{\mu 6} \overline{F}_{\nu 6} = e^2 \Big\{ & \tfrac{1}{4} s^{-6} \epsilon^{-2} (\ddot{z}^2 \dot{z}_\mu \dot{z}_\nu - c^2 \ddot{z}_\mu \ddot{z}_\nu) \\ & + (s^{-2} \xi^{-6} + \tfrac{1}{4} s^{-6} \epsilon^{-4} \ddot{z}^2)(\epsilon^2 \dot{z}_\mu \dot{z}_\nu - c^2 \xi_\mu \xi_\nu) \\ & - \tfrac{1}{2} s^{-4} \epsilon^{-4} (2K \dot{z}_\mu \dot{z}_\nu - c^2 \ddot{z}_\mu \xi_\nu - c^2 \ddot{z}_\nu \xi_\mu) \\ & - \tfrac{1}{2} s^{-4} \epsilon^{-4} (\epsilon^2 \dddot{z}_\mu \dot{z}_\nu + \epsilon^2 \dddot{z}_\nu \dot{z}_\mu - \xi_6 \dddot{z}_6 \xi_\mu \dot{z}_\nu - \xi_6 \dddot{z}_6 \xi_\nu \dot{z}_\mu \\ & - 2 \dot{z}^2 \xi_\mu \xi_\nu) \Big\} \end{aligned} \qquad (3.112)$$

to the -4th, - 3rd, and - 2nd orders,

$$\begin{aligned} \overline{F}_{\mu 6} \overline{F}_{\nu 6} \, d\sigma_\nu = e^2 \Big\{ & \tfrac{1}{4} c^{-1} s^{-4} \epsilon^{-1} (-c^2 K \ddot{z}_\mu) + (c^{-1} \epsilon^{-5} + \tfrac{1}{4} c^{-1} s^{-4} \epsilon^{-3} \ddot{z}^2)(-c^2 \epsilon^2 \xi_\mu) \\ & - \tfrac{1}{2} c^{-1} s^{-2} \epsilon^{-3} (-c^2 \epsilon^2 \ddot{z}_\mu - c^2 K \xi_\mu) - \tfrac{1}{2} c^{-1} s^{-2} \epsilon^{-3} (-2\epsilon^2 \ddot{z}^2 \xi_\mu) \Big\} \, d\Omega \, d\tau \end{aligned}$$

$$\bar{F}_{\mu\sigma}\bar{F}_{\nu\sigma}\,d\sigma_\nu = e^2\Big\{\tfrac{1}{4}cs^{-4}\epsilon^{-1}(2s^2-\kappa)\ddot{z}_\mu - \tfrac{1}{2}cs^{-2}\epsilon^{-3}(2s^2-\kappa)\xi_\mu$$
$$-\tfrac{1}{4}c^{-1}s^{-4}\epsilon^{-1}(c^2+4s^2)\dot{z}^2\xi_\mu\Big\}\,d\Omega\,d\tau$$

(3.113)

to the - 2nd, - 1st, and 0th orders, and

$$\bar{F}_{\sigma\tau}\bar{F}_{\sigma\tau}\,d\sigma_\mu = e^2c^{-1}s^2\epsilon\Big\{\tfrac{1}{4}s^{-6}\epsilon^{-2}(-2c^2\dot{z}^2) + \big(s^{-2}\epsilon^{-6}+\tfrac{1}{4}s^{-6}\epsilon^{-4}\dot{z}^2\big)(-2c^2\epsilon^2)$$
$$-\tfrac{1}{2}s^{-4}\epsilon^{-4}(-4c^2\kappa) - \tfrac{1}{2}s^{-4}\epsilon^{-4}(-4\epsilon^2\dot{z}^2)\Big\}\xi_\mu\,d\Omega\,d\tau$$
$$= e^2\Big\{-2cs^{-2}\epsilon^{-3}(s^2-\kappa)\xi_\mu - c^{-1}s^{-4}\epsilon^{-1}(c^2+2s^2)\dot{z}^2\xi_\mu\Big\}\,d\Omega\,d\tau$$

(3.114)

to the - 2nd, - 1st, and 0th orders. Using the expansion

$$s^{-4}(2s^2-\kappa) = 2c^{-2}\,\frac{1+\frac{\kappa}{2c^2}}{\left(1+\frac{\kappa}{c^2}\right)^2} = 2c^{-2}\left(1-\tfrac{3}{2}c^{-2}\kappa+\cdots\right) \qquad (3.115)$$

we finnaly have, correct to the - 2nd, - 1st, and 0th orders,

$$\bar{F}_{\mu\sigma}\bar{F}_{\nu\sigma}\,d\sigma_\nu - \tfrac{1}{4}\bar{F}_{\sigma\tau}\bar{F}_{\sigma\tau}\,d\sigma_\mu$$
$$= e^2\Big\{\tfrac{1}{4}cs^{-4}\epsilon^{-1}(2s^2-\kappa)\ddot{z}_\mu - \tfrac{1}{2}c\epsilon^{-3}\xi_\mu + \tfrac{1}{2}c^{-1}s^{-2}\epsilon^{-1}\dot{z}^2\xi_\mu\Big\}\,d\Omega\,d\tau$$
$$= e^2\Big\{\tfrac{1}{2}c^{-1}\epsilon^{-1}\ddot{z}_\mu - \tfrac{3}{4}c^{-3}\epsilon^{-1}\kappa\ddot{z}_\mu - \tfrac{1}{2}c\epsilon^{-3}\xi_\mu + \tfrac{1}{2}c^{-3}\epsilon^{-1}\dot{z}^2\xi_\mu\Big\}\,d\Omega\,d\tau$$

(3.116)

Next, we compute

$$\bar{F}^{free}_{\mu\sigma}\bar{F}_{\nu\sigma}\,d\sigma_\nu = e\,c^{-1}s\epsilon^{-2}\bar{F}^{free}_{\mu\sigma}\big(\dot{\xi}_\nu\xi_\sigma - \dot{z}_\sigma\xi_\nu\big)\xi_\nu\,d\Omega\,d\tau$$
$$= -e\,\bar{F}^{free}_{\mu\sigma}\dot{z}_\sigma\,d\Omega\,d\tau \qquad (3.117)$$

and

$$\bar{F}_{\mu\sigma}\bar{F}^{free}_{\nu\sigma}\,d\sigma_\nu - \frac{1}{2}\bar{F}^{free}_{\sigma\tau}\bar{F}_{\sigma\tau}\,d\sigma_\mu = \bar{F}^{free}_{\sigma\tau}\left(\bar{F}_{\mu\tau}\,d\sigma_\sigma - \frac{1}{2}\bar{F}_{\sigma\tau}\,d\sigma_\mu\right)$$

$$= e\,c^{-1}\,\epsilon^{-2}\,\bar{F}^{free}_{\sigma\tau}\left[\left(\dot{z}_\mu\xi_\tau - \dot{z}_\tau\xi_\mu\right)\xi_\sigma - \frac{1}{2}\left(z_\sigma\xi_\tau - \dot{z}_\tau\xi_\sigma\right)\xi_\mu\right]d\Omega\,d\tau$$

$$= 0 \qquad (3.118)$$

correct to the 0th order.

We are now readily to evaluate the integral (3.105). We shall carry out the integration over the solid angle first. Using the fact that, by symetry,

$$\frac{1}{4\pi}\int_{4\pi}\xi_\mu\,d\Omega = 0 \quad , \quad \frac{1}{4\pi}\int_{4\pi}\kappa\,d\Omega = 0 \qquad (3.119)$$

and taking ϵ so small that $\bar{F}^{free}_{\mu\nu}$ may be regarded as constant and equal to its value at $z_\mu(\tau)$, in the integration, we obtain

$$\frac{1}{c}\int_{4\pi}T_{\mu\nu}\,d\sigma_\nu = \int\left\{\frac{1}{2}\,e^2c^{-2}\epsilon^{-1}\ddot{z}_\mu - e\,c^{-1}\bar{F}^{free}_{\mu\sigma}(z)\,\dot{z}_\sigma\right\}d\tau \qquad (3.120)$$

Now, let $\Delta\sigma_1$ and $\Delta\sigma_2$ denote the perpendicular cross sections of the tube at the proper time τ_1 and τ_2 respectively. Then using Gauss's theorem and the conservation equation (3.31), we may write, if $\tau_2 > \tau_1$

$$\frac{1}{c}\int_{\tau_1}^{\tau_2}T_{\mu\nu}\,d\sigma_\nu + \frac{1}{c}\int_{\Delta\sigma_2}T_{\mu\nu}\,d\sigma_\nu - \frac{1}{c}\int_{\Delta\sigma_1}T_{\mu\nu}\,d\sigma_\nu = \int_{\Delta V}T_{\mu\nu,\nu}\,dx = 0 \qquad (3.121)$$

where ΔV denotes the space-time region enclosed by $\Delta\sigma_1$, $\Delta\sigma_2$ and the tube. In the integrations over $\Delta\sigma_1$ and $\Delta\sigma_2$, the directed

surface element $d\sigma_\nu$ has in both cases a negative time-like orientation in order that Gauss's theorem be applied correctly to a non-positive-definite space. In the limit $\epsilon \to 0$ only the particle stress will contribute to the integrals over $\Delta\sigma_1$ and $\Delta\sigma_2$. Hence, using (3.24), (3.26) and (3.120), we obtain

$$\int_{\tau_1}^{\tau_2} \left\{ \tfrac{1}{2} e^2 c^{-2} \epsilon^{-1} \ddot{z}_\mu - e c^{-1} \bar{F}^{ext}_{\mu\nu}(z)\, \dot{z}_\nu \right\} d\tau + m \dot{z}_\mu(\tau_2) - m \dot{z}_\mu(\tau_1) = 0 \,. \tag{3.122}$$

Taking τ_2 infinitesimally close to τ_1, we obtain the differential equation

$$\left(m + \tfrac{1}{2} e^2 c^{-2} \epsilon^{-1} \right) \ddot{z}_\mu = \frac{e}{c} \bar{F}^{ext}_{\mu\nu}(z)\, \dot{z}_\nu \tag{3.123}$$

which is reminiscent of Eq. (3.15). Eq. (3.123) can, in fact, be derived directly from (3.15) by making use of (3.103). Since $\zeta_\mu \to 0$, $K \to 0$, and $S \to c$ as $\epsilon \to 0$ we have, from (3.103)

$$\bar{F}_{\mu\nu}(z) = -\tfrac{1}{2} e c^{-3} \epsilon^{-1} \left(\dot{z}_\mu \ddot{z}_\nu - \dot{z}_\nu \ddot{z}_\mu \right) \tag{3.124}$$

and

$$\bar{F}_{\mu\nu}(z)\, \dot{z}_\nu = -\tfrac{1}{2} e c^{-1} \epsilon^{-1} \ddot{z}_\mu \tag{3.125}$$

Substitution into (3.15), with $F_{\mu\nu} = \bar{F}_{\mu\nu} + \bar{F}^{ext}_{\mu\nu}$ gives (3.123).

The important difference between equations (3.15) and (3.123) is that the divergence which is inherent in them has, in the latter equation, been completely isolated and explicitely exhibited. The divergent term is seen to have precisely the character of a self-energy term, the quantity $\frac{1}{2} e^2 c^{-2} \epsilon^{-1}$ being the mass correction. We may effect a classical "mass renormalization" by introducing the experimentally observed mass.

$$m_{obs} = m + \lim_{\epsilon \to 0} \tfrac{1}{2} \frac{e^2}{c^2 \epsilon} \,, \tag{3.126}$$

and rewriting (3.123) in the form

$$m_{obs}\,\ddot{z}_\mu = \frac{e}{c}\,\bar{F}^{ret}_{\mu\nu}(z)\,\dot{z}_\nu \tag{3.127}$$

All quantities appearing in (3.127) are now finite.

For practical applications it is useful to rewrite (3.127) in a form which recognizes the usual boundary conditions which we impose. In practise only the incoming field $F^{in}_{\mu\nu}$ is known in advance. Therefore, remembering $\bar{F}^{ret}_{\mu\nu} = F^{in}_{\mu\nu} + \frac{1}{2} F^{rad}_{\mu\nu}$ and making use of (3.104), we write

$$m_{obs}\,\ddot{z}_\mu = \frac{e}{c}\,F^{in}_{\mu\nu}(z)\,\dot{z}_\nu + \frac{e}{2c}\,F^{rad}_{\mu\nu}(z)\,\dot{z}_\nu$$

$$= \frac{e}{c}\,F^{in}_{\mu\nu}(z)\,\dot{z}_\nu - \frac{2}{3}\frac{e^2}{c^3}\left(\ddot{z}^2\dot{z}_\mu - c^2\dddot{z}_\mu\right) \tag{3.128}$$

$F^{rad}_{\mu\nu}$ is now seen to have the interpretation as the field which describes the force of radiative reaction on the particle.

Run-away solutions

Equations (3.128) have some peculiar properties. Let us consider the case in which $F^{in}_{\mu\nu} = 0$, and let us suppose that the motion of the particle takes place in a straight line parallel to the x - axis so that

$$\dot{y} = 0 \qquad \text{and} \qquad \dot{z} = 0 \tag{3.129}$$

Equations (3.128) then reduce to

$$\gamma\ddot{x} - \dddot{x} + \frac{1}{c^2}\left(\ddot{x}^2 - c^2\ddot{t}^2\right)\dot{x} = 0 \tag{3.130}$$

$$\gamma\ddot{t} - \dddot{t} + \frac{1}{c^2}\left(\ddot{x}^2 - c^2\ddot{t}^2\right)\dot{t} = 0 \tag{3.131}$$

where

$$\gamma = \frac{3}{2}\,c/r_0 \quad , \quad r_0 = \frac{e^2}{m_{obs}\,c^2} \tag{3.132}$$

Eq. (3.131) is actually implied by Eq. (3.130) through the relation

$$\dot{x}^2 - c^2\dot{t}^2 = -c^2 \tag{3.133}$$

We have

$$c\dot{t} = \sqrt{c^2+\dot{x}^2} \quad , \quad c\ddot{t} = \frac{\dot{x}\ddot{x}}{\sqrt{c^2+\dot{x}^2}} \tag{3.134}$$

and

$$0 = r\ddot{x} - \dddot{x} + \frac{1}{c^2}\left(\ddot{x}^2 - \frac{\dot{x}^2\ddot{x}^2}{c^2+\dot{x}^2}\right)\dot{x}$$

$$= r\ddot{x} - \dddot{x} + \frac{\dot{x}\ddot{x}^2}{c^2+\dot{x}^2} \tag{3.135}$$

The physically sensible solution of this equation is simply

$$\ddot{x} = 0, \quad \dot{x} = \text{constant}, \quad \dot{t} = \text{constant} \tag{3.136}$$

corresponding to uniform motion. However, Eq. (3.135) also allows of "self-accelerated" motion. For, suppose $\ddot{x} \neq 0$. Then, dividing through by $\ddot{x}$ we obtain

$$r - \frac{\dddot{x}}{\ddot{x}} + \frac{\dot{x}\ddot{x}}{c^2+\dot{x}^2} = 0 \tag{3.137}$$

which, on integration, gives

$$r\tau - \log \ddot{x} + \tfrac{1}{2}\log(c^2+\dot{x}^2) = A \tag{3.138}$$

If we choose the origin of τ so that $A = \log r$, this becomes

$$\frac{\ddot{x}}{\sqrt{c^2+\dot{x}^2}} = r e^{r\tau} \tag{3.139}$$

which yields

$$\sinh^{-1}\frac{\dot{x}}{c} = e^{r\tau} + B \tag{3.140}$$

or

$$\dot{x} = c \sinh (e^{\gamma\tau} + B), \tag{3.141}$$

$$\dot{t} = \cosh (e^{\gamma\tau} + B). \tag{3.142}$$

Solutions (3.141, 142) are known as "runaway" solutions. According to these solutions the velocity of the particle, starting out with the value $c \tanh B$, builds up asymptotically to the velocity of light in an incredibly short time (i.e. of the order of time it takes for a pulse of light to cross a distance z_0.). Such non-physical solutions must be expressly forbidden.

Non covariant form of the equation

It will be convenient to rewrite equations (3.128) in non-covariant form. If the position vector of the particle be denoted by r then the spatial part of equations (3.128) takes the form

$$m \ddot{r} - \frac{2}{3}\frac{e^2}{c^3}\dddot{r} + \frac{2}{3}\frac{e^2}{c^5}\left(\ddot{r}^2 - c^2\ddot{t}^2\right)\dot{r} = \frac{e}{c}\dot{r} \times H^{in}(r,t) + e\dot{t}\, E^{in}(r,t) \tag{3.143}$$

where use has been made of the explicit form (3.8) for $F^{in}_{\mu\nu}$, and the subscript "obs" has been dropped from m, it being henceforth understood that it is the experimentally observed mass which is involved. The proper time derivations in (3.143) may be replaced by ordinary time derivations through use of the relation

$$-c^2 = \dot{r}^2 - c^2\dot{t}^2 = \dot{t}^2\left(v^2 - c^2\right) \tag{3.14'}$$

where

$$v = \frac{dr}{dt} \tag{3.145}$$

We then have

$$\dot{r} = \dot{t}\frac{dr}{dt} = \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.146}$$

$$\ddot{r} = \dot{t}\frac{d}{dt}\left(\dot{t}\frac{dr}{dt}\right) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\ \frac{d}{dt}\ \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3.147}$$

Also

$$c\dot{t} = \sqrt{c^2 + \dot{r}^2} \quad , \quad c\ddot{t} = \frac{\dot{r}\ddot{r}}{\sqrt{c^2 + \dot{r}^2}} \tag{3.148}$$

$$\ddot{r}^2 - c\ddot{t}^2 = \ddot{r}^2 - \frac{(\dot{r},\ddot{r})^2}{c^2 + \dot{r}^2} = \frac{c^2\ddot{r}^2 + (\dot{r}\times\ddot{r})^2}{c^2\dot{t}^2} \tag{3.149}$$

Hence, dividing Eq. (3.143) by $\dot{t}$, we obtain

$$m\frac{d}{dt}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}} - \frac{2}{3}\frac{e^2}{c^3}\frac{d}{dt}\left[\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{d}{dt}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}\right]$$

$$+\frac{2}{3}\frac{e^2}{c^5}\left[\left(\frac{d}{dt}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}\right)^2 + \frac{1}{c^2}\left(\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}\times\frac{d}{dt}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}\right)^2\right]v$$

$$= \frac{e}{c}v\times H^{in}(r,t) + eE^{in}(r,t) . \tag{3.150}$$

Example

Suppose the incoming field $F^{in}_{\mu\nu}$ is that of a plane wave of angular frequency ω traveling in the z direction and polarized in the x direction. Then the 4-vector potential has the form

$$(A^{in}_{\mu}) = \left(\mathcal{A}\sin\frac{\omega}{c}(z-ct), 0, 0, 0\right), \tag{3.151}$$

the electric and magnetic vectors being given by

$$E^{in} = \left(\frac{\omega}{c}\mathcal{A}\cos\frac{\omega}{c}(z-ct), 0, 0\right) \tag{3.152}$$

$$H^{in} = \left(0, \frac{\omega}{c}\mathcal{A}\cos\frac{\omega}{c}(z-ct), 0\right) \tag{3.153}$$

We have here denoted the amplitude of the vector potential by $\mathcal{A}$,

because we shall suppose this amplitude to be infinitesimally small, corresponding to the limit of an infinitely weak wave. Eqs. (3.150) are consistent with

$$y = \text{constant} \quad , \quad \frac{dy}{dt} = 0 \tag{3.154}$$

regardless of the strenght of the incoming wave. For the motions of the x and z coordinates, however, we have, correct to the second infinitesiaml order, supposing that the average position of the charged particle is initially ($t = 0$) at rest at the origin,

$$m \frac{d^2x}{dt^2} - \frac{2}{3} \frac{e^2}{c^3} \frac{d^3x}{dt^3} = e \frac{\omega}{c} \delta A \cos \omega t \tag{3.155}$$

$$m \frac{d^2z}{dt^2} - \frac{2}{3} \frac{e^2}{c^3} \frac{d^3z}{dt^3} = \frac{e}{c} \frac{dx}{dt} \frac{\omega}{c} \delta A \cos \omega t \tag{3.156}$$

It is evident that x is of the first infinitesimal order and z of the second.

The solution of (3.155) which corresponds to an average position at rest must have the form

$$x = a_1 \cos \omega t + a_2 \sin \omega t \tag{3.157}$$

Substituting (3.157) into (3.155), we obtain

$$-a_1 \omega^2 \cos \omega t - a_2 \omega_2 \sin \omega t - \frac{1}{\tau} \left(a_1 \omega^3 \sin \omega t - a_2 \omega^3 \cos \omega t \right) = \frac{e}{m} \frac{\omega}{c} \delta A \cos \omega t \tag{3.158}$$

which implies

$$-a_1 \omega^2 + a_2 \frac{\omega^3}{\tau} = \frac{e}{m} \frac{\omega}{c} \delta A \tag{3.159}$$

$$-a_1 \frac{\omega^3}{\alpha} - a_2 \omega^2 = 0 \tag{3.160}$$

The solution of these simultaneous equations is

$$a_1 = - \frac{\frac{e}{m} \frac{1}{\omega c}}{1 + \frac{\omega^2}{\tau}} \delta A \tag{3.161}$$

$$a_2 = \frac{\frac{e}{mc}\frac{1}{r^2}}{1+\frac{\omega^2}{r^2}}\,\delta A \tag{3.162}$$

so that

$$\frac{dx}{dt} = \frac{\frac{e}{mc}}{1+\frac{\omega^2}{r^2}}\,\delta A \sin\omega t + \frac{\omega}{r}\,\frac{\frac{e}{mc}}{1+\frac{\omega^2}{r^2}}\,\delta A \cos\omega t \tag{3.163}$$

Substitution of (3.163) into (3.156) gives

$$\frac{d^2z}{dt^2} - \frac{1}{r}\frac{d^3z}{dt^3} = \frac{\left(\frac{e}{mc}\right)^2}{1+\frac{\omega^2}{r^2}}\,\frac{\omega}{c}\,(\delta A)^2\left\{\frac{1}{2}\sin 2\omega t + \frac{1}{2}\frac{\omega}{r}\left(1+\cos 2\omega t\right)\right\} \tag{3.164}$$

The solution of (3.164) which corresponds to an average position initially at rest must have the form

$$z = a_3\cos 2\omega t + a_4 \sin 2\omega t + \frac{1}{2}a_5 t^2 \tag{3.165}$$

Substitution of (3.165) into (3.164) leads to the equations

$$-a_3(2\omega)^2 + a_4\frac{(2\omega)^3}{r} = \frac{1}{2}\,\frac{\left(\frac{e}{mc}\right)^2}{1+\frac{\omega^2}{r^2}}\,\frac{\omega}{c}\,\frac{\omega}{r}\,(\delta A)^2 \tag{3.166}$$

$$-a_3\frac{(2\omega)^3}{r} - a_4(2\omega)^2 = \frac{1}{2}\,\frac{\frac{e}{mc}}{1+\frac{\omega^2}{r^2}}\,\frac{\omega}{c}\,(\delta A)^2 \tag{3.167}$$

$$a_5 = \frac{1}{2}\,\frac{\left(\frac{e}{mc}\right)^2}{1+\frac{\omega^2}{r^2}}\,\frac{\omega}{c}\,\frac{\omega}{r}\,(\delta A)^2 \tag{3.168}$$

The solution of Eqs. (3.166, 167) is

$$a_3 = -\frac{3}{8}\,\frac{\left(\frac{e}{mc}\right)^2}{\left(1+\frac{\omega^2}{r^2}\right)\left(1+4\frac{\omega^2}{r^2}\right)}\,\frac{1}{rc}\,(\delta A)^2 \tag{3.169}$$

$$a_4 = -\frac{1}{8}\,\frac{\left(\frac{e}{mc}\right)^2\left(1-2\frac{\omega^2}{r^2}\right)}{\left(1+\frac{\omega^2}{r^2}\right)\left(1+4\frac{\omega^2}{r^2}\right)}\,\frac{1}{\omega c}\,(\delta A)^2 \tag{3.170}$$

The mean position of the particle is seen to suffer a constant acceleration in the Z direction. Explicitely we write

$$\bar{z} = \frac{1}{2} a_5 t^2 \tag{3.171}$$

The mean momentum at time t is therefore

$$m \frac{d\bar{z}}{dt} = m a_5 t \tag{3.172}$$

That is, the mean particle momentum increases linearly with time. We shall now show that this increase is caused by the scattering by the particle of energy and momentum out of the incident electromagnetic wave.

Thomson Scattering

The energy in the incident field is given (see Eqs. (3.22, 23, 29) for definitions) by

$$E^{in} = -ic\, p_4^{in} = -i \int_\sigma T_{4\nu}^{in}\, d\sigma_\nu = \int_\infty U^{in}\, d_3\mathbf{r} \tag{3.173}$$

where U^{in} is the energy density :

$$\begin{aligned} U_{in} &= -T_{44}^{in} \\ &= -\frac{1}{4\pi}\left(F_{46}^{in} F_{46}^{in} - \frac{1}{4} F_{6\tau}^{in} F_{6\tau}^{in} \right) \\ &= -\frac{1}{4\pi}\left(-E^{in\,2} - \frac{1}{4}\left(2H^{in\,2} - 2E^{in\,2} \right) \right) \\ &= \frac{1}{8\pi}\left(E^{in\,2} + H^{in\,2} \right) \\ &= \frac{1}{4\pi} \frac{\omega^2}{c^2} \left(\delta A \right)^2 \cos^2 \frac{\omega}{c} \left(z - ct \right) \end{aligned} \tag{3.174}$$

The tensor $T_{\mu\nu}^{in}$ satisfies the conservation equation

$$0 = -ic\, T_{4\nu,\nu}^{in} = \frac{\partial U^{in}}{\partial t} + \nabla \cdot S^{in} \tag{3.175}$$

where S^{in} is the energy flux vector :

$$S^{in} = \left(-ic\, T^{in}_{4i}\right) = -\frac{ic}{4\pi}\left(F^{in}_{4\sigma}\, F^{in}_{i\sigma}\right) = \frac{c}{4\pi}\left(E^{in}_{j} F^{in}_{ij}\right)$$

$$= \frac{c}{4\pi} E^{in} \times H^{in} = \frac{c}{4\pi}\left(0, 0, \frac{\omega^2}{c^2}(\sqrt{A})^2 \cos^2\frac{\omega}{c}(z-ct)\right) \quad (3.176)$$

The non-vanishing z -component of the energy flux vector is seen to be simply C times the energy density.

Now let dE^{rad}/dt be the rate of energy radiation by the particle. Then the cross section for scattering of the electromagnetic field by the particle is defined by

$$\sigma = \frac{\overline{\frac{dE}{dt}}^{rad}}{\overline{S}^{in}_{3}} \quad (3.176')$$

where the bars denote time averages, The rate of energy radiation is given by the radiation reaction terms in the equations of motion (3.128). Referring to Eq. (3.19) we see that we must have

$$\frac{dE^{rad}}{d\tau} = -mc\,\ddot{z}^{rad}_{0} = \frac{2}{3}\frac{e^2}{c^4}\left(\ddot{z}_0^2 \dot{z}_0 - c^2 \dddot{z}_0\right) \quad (3.176'')$$

Since the velocity of the particle remains infinitesimally small in the present case we have $\dot{z}_0 = c$, and we may replace τ by t writing Eq. (3.176) in the form

$$\frac{dE^{rad}}{dt} = \frac{2}{3}\frac{e^2}{c^3}\left(\frac{d^2x}{dt^2}\right)^2$$

$$= \frac{2}{3}\frac{e^2}{c^3}\omega^4\left(a_1 \cos\omega t + a_2 \sin\omega t\right)^2 \quad (3.177)$$

Taking the mean value, we have

$$\frac{\overline{dE}^{rad}}{dt} = \frac{1}{3}\frac{e^2}{c^3}\omega^4\left(a_1^2 + a_2^2\right)$$

$$\overline{\frac{dE}{dt}}^{rad} = \frac{1}{3}\,\frac{e^4}{m^2c^5}\,(\delta A)^2\,\frac{\omega^2}{1+\frac{\omega^2}{\gamma^2}} \tag{3.178}$$

Since

$$\bar{S}_3^{in} = \frac{1}{8\pi}\,\frac{\omega^2}{c}\,(\delta A)^2 \tag{3.179}$$

we obtain for the scattering cross section

$$\sigma = \frac{8\pi}{3}\,\frac{e^4}{m^2c^4}\,\frac{1}{1+\frac{\omega^2}{\gamma^2}} = \frac{8}{3}\,\pi\,r_0^2\,\frac{1}{1+\frac{\omega^2}{\gamma^2}} \tag{3.180}$$

This is the Thomson scattering formula with a "damping factor" $\left(1+\frac{\omega^2}{\gamma^2}\right)^{-1}$ which reduces the cross section at high frequencies.

Let us now look at the momentum in the incident field.

We have

$$P^{in} = \frac{1}{c}\int_\sigma \left(T_{i\nu}^{in}\right) d\sigma_\nu = -\frac{i}{c}\int_{-\infty} \left(T_{i4}^{in}\right) d_3 r = \int_{-\infty} G^{in}\, d_3 r \tag{3.181}$$

where G^{in} is the momentum density :

$$G^{in} = \frac{1}{c^2}\,S^{in} \tag{3.182}$$

We again have a conservation equation

$$0 = \left(T_{i\nu,\nu}^{in}\right) = \frac{\partial G^{in}}{\partial t} + \nabla\,\Theta^{in} \tag{3.183}$$

where Θ^{in} is the momentum flux dyadic :

$$\Theta^{in} = \left(T_{ij}^{in}\right) \tag{3.184}$$

In the present example the $\vec{z}$ -component of the momentum flux is of interest.

$$\left(\Theta_{3j}\right) = \left(T_{3j}^{in}\right) = \frac{1}{4\pi}\left(F_{3\sigma}^{in}\,F_{j\sigma}^{in} - \frac{1}{4}\,\delta_{3j}\,F_{c\tau}^{in}\,F_{\sigma\tau}^{in}\right)$$

$$(\Theta_{3j}) = \frac{1}{4\pi}\left(0, 0, H_2^{in^2} - \frac{1}{4}\left(2H_2^{in^2} - 2E_1^{in^2}\right)\right)$$

$$= \left(0, 0, U^{in}\right) \tag{3.185}$$

Now, the mean rate of increase of momentum of the particle is given by

$$m\frac{d^2\bar{z}}{dt^2} = ma_5 = \frac{1}{3}\frac{\omega^2}{c^2}\frac{e^4}{m^2c^4}\frac{(\sqrt{A})^2}{1+\frac{\omega^2}{\gamma^2}} \tag{3.186}$$

in which we have used (3.168) and the fact that

$$\frac{1}{\tau} = \frac{2}{3}\frac{e^2}{mc^3} \tag{3.187}$$

If the radiation scattered by the particle is scattered symmetrically fore and aft, so that on the average the particle recoil is just what it would be if the particle absorbed all the momentum scattered out of the field, then the mean rate of increase of the particle momentum should be given by

$$\sigma\,\overline{U^{in}} = \frac{1}{3}\frac{\omega^2}{c^2}\frac{e^4}{m^2c^4}\frac{(\sqrt{A})^2}{1+\frac{\omega^2}{\gamma^2}} \tag{3.188}$$

Comparison shows, in fact, that expressions (3.186) and (3.188) are identical.

Problem XXV: Show that the scattering is indeed symetrical fore and aft. In particular, show thaht the differential scattering cross section is

$$d\sigma = \frac{e^4}{m^2c^4}\frac{\sin^2\beta}{1+\frac{\omega^2}{\gamma^2}}\,d\Omega$$

where β is the angle between the scattering direction and the

z -axis. By averaging over the different polarizations of the incident field show therefore that the differetnial scattering cross section for unpolarized monochromatic radiation is

$$d\sigma = \frac{e^4}{m^2c^4} \frac{\frac{1}{2} + \frac{1}{2}\cos^2\theta}{1 + \frac{\omega^2}{\gamma^2}} d\Omega$$

where θ is the angle between the scattering direction and the z -axis. Use may conveniently be made of the Lienard-Wiechert potentials obtained in Problem XXIV.

Several interacting charged particles

The classical mass renormalization carried out in the preceding pages may readily be extended to the case of several interacting charged particles. Each particle will have associated with it its own retarded and advanced fields, $F^{ret}_{n\nu\mu}$ and $F^{ret}_{n\nu\mu}$ respectively. (The subscript n labels the particle in question.) In terms of the incident and radiation fields associated with the n th particle, namely,

$$F^{in}_{n\mu\nu} = F_{\mu\nu} - F^{ret}_{n\mu\nu}, \tag{3.189}$$

$$F^{rad}_{n\mu\nu} = F^{ret}_{n\mu\nu} - F^{adv}_{n\mu\nu}, \tag{3.190}$$

we have, as before,

$$m_n \ddot{z}_{n\mu} = \frac{e_n}{c} F^{in}_{n\mu\nu}(z_n)\dot{z}_{n\nu} + \frac{e_n}{2c} F^{rad}_{n\mu\nu}(z_n)\dot{z}_{n\nu}$$

$$= \frac{e_n}{c} F^{in}_{n\mu\nu}(z_n)\dot{z}_{n\nu} - \frac{2}{3}\frac{e_n^2}{c^5}\left(\ddot{z}_n^2\dot{z}_{n\mu} - c^2\dddot{z}_{n\mu}\right), \tag{3.191}$$

there being one such equation for each particle. Attention should be called to the fact that each particle will have its own proper time, and the dots above naturally denote differentiation with respect to the proper time of the particle in question.

In practise, the individual incident fields $F^{in}_{n\mu\nu}$ are not specified. Rather it is the field incident on the entire assembly

of particles which is known in advance. This field is defined by

$$F^{in}_{\mu\nu} = F_{\mu\nu} - \sum_m F^{ret}_{m\mu\nu}$$

$$= F^{in}_{n\mu\nu} - \sum_{\substack{m \\ m\neq n}} F^{ret}_{m\mu\nu} \qquad (3.192)$$

In terms of this field the equations of motion become

$$m \ddot{z}_{n\mu} + \frac{2}{3}\frac{e_n^2}{c^5}\left(\dot{z}^2 \dot{z}_{n\mu} - c^2 \dddot{z}_{n\mu}\right)$$

$$= \frac{e_n}{c} F^{in}_{\mu\nu}(z_n)\, \dot{z}_{n\nu} + \frac{e_n}{c} \sum_{\substack{m \\ m\neq n}} F^{ret}_{m\mu\nu}(z_n)\, \dot{z}_{n\nu} \qquad (3.193)$$

Bibliography

P.A.M. Dirac, Proceedings of the Royal Society, A 167 (1938), p. 148

COURS DE L'ÉCOLE D'ÉTÉ DE PHYSIQUE THÉORIQUE

QUANTUM MECHANICS

Bryce S. DE WITT

Radiation Laboratory
University of California

★ ★ ★

LES HOUCHES
ÉTÉ 1953

Q U A N T U M M E C H A N I C S

Lecture Notes by
Bryce S. De Witt
Radiation Laboratory
University of California

Chapters IV, V, VI, VII.

Université de Grenoble
Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

4.- THE CLASSICAL THEORY OF SPINNING PARTICLES.

Consider an assembly of point particles of masses m_n interacting with one another through instantaneous forces describable by a potential function $V(r_1, \ldots)$ where r_n denotes the position of the n th particle. The Lagrangian function for the system is

$$L = \sum_n \frac{1}{2} m_n \dot{r}_n^2 - V \qquad (4.1)$$

and the Hamiltonian function is

$$H = \sum_n \frac{1}{2m_n} p_n^2 + V, \qquad (4.2)$$

the canonical momenta being

$$p_n = \frac{\partial L}{\partial \dot{r}_n} = m_n \dot{r}_n \quad , NS. \qquad (4.3)$$

Let the entire assembly of particles be rotated through an infinitesimal angle $\delta\alpha$ about a certain axis through the origin. The direction of the axis may be specified by allowing $\delta\alpha$ to be a vector quantity. The variations in the coordinates and momenta due to this rotation are given by

$$\delta r_n = \delta\alpha \times r_n \quad , \quad \delta p_n = \delta\alpha \times p_n \, . \qquad (4.4)$$

Instead of regarding this variation in the canonical variables as being due to a rotation of the mechanical system, we may equally well regard it as being due to a rotation of the coordinate system through an angle $-\delta\alpha$. Such a rotation is an orthogonal transformation of the coordinates, and, in fact, it may readily be shown that the transformation (4.4) is simply a point transformation. The canonicity of (4.4) may be proved directly by calculating the Poisson brackets of the new variables. We shall, however, do it by finding a

generating function for the transformation.

In component form Eq. (4.4) becomes

$$\delta r_{ni} = \epsilon_{ijk}\,\delta\alpha_j\, r_{nk} \quad , \quad \delta p_{ni} = \epsilon_{ijk}\,\delta\alpha_j\, p_{nk} \,. \tag{4.5}$$

But for canonicity we must have

$$\delta r_{ni} = \frac{\partial S}{\partial p_{ni}} \quad , \quad \delta p_{ni} = -\frac{\partial S}{\partial r_{ni}} \tag{4.6}$$

for some function S. Equations (4.5) and (4.6) may immediatly be integrated to give

$$S = \sum_n \epsilon_{ijk}\,\delta\alpha_i\, r_{nj}\, p_{nk} = J\cdot\delta\alpha \tag{4.7}$$

where

$$J = \sum_n r_n \times p_n \tag{4.8}$$

J is well known as the angular momentum of the assembly about the origin. It is evident that equations (4.6) may be generalized to

$$\delta A = (A, S) = (A, J_i)\,\delta\alpha_i \tag{4.9}$$

where A is any vector or tensor quantity composed out of the canonical variables and δA denotes its variation under the operation $\delta\alpha$. In particular, since $\delta J_i = \epsilon_{ijk}\,\delta\alpha_j J_k$, we have

$$\epsilon_{ijk}\,\delta\alpha_j J_k = (J_i, J_s)\,\delta\alpha_j \tag{4.10}$$

Since $\delta\alpha$ is arbitrary this implies the following important Poisson bracket relation :

$$(J_i, J_j) = \epsilon_{ijk} J_k \,. \tag{4.11}$$

Suppose now that the form of the Hamiltonian function (4.2) remains invariant under the canonical transformation (4.5). (This will be the case, for example, if V depends only on the distances between the particles, the angles between the straight lines joining them, and the components of the position vectors r_n parallel to $\delta\alpha$.) Then the results of chapter I, Eq. (1.69), tell us that the function S of (4.7) is a constant of the motion. This is the law of conservation of angular momentum about a fixed axis.

Suppose that the form of the Hamiltonian function remains invariant under rotations $\delta\alpha$, $\delta\beta$ about two distinct axes. (This will be the case if V depends only on the distances $|r_m - r_n|$ between the particles and on the angles between the straight lines joining them.) Then the quantities $J\cdot\delta\alpha$ and $J\cdot\delta\beta$ will both be constants of the motion, and the results of chapter I, page 18, tell us the quantity

$$(J\cdot\delta\alpha, J\cdot\delta\beta) = J\cdot\delta\gamma \tag{4.12}$$

with

$$\delta\gamma = \delta\alpha \times \delta\beta, \tag{4.13}$$

is also a constant of the motion. Since $\delta\alpha$ and $\delta\beta$ are not parallel, the vectors $\delta\alpha$, $\delta\beta$, $\delta\gamma$ are linearly independent, and hence the components of J are, independently, constants of the motion. That is,

$$\dot{J} = 0 \tag{4.14}$$

This is the law of conservation of angular momentum about an arbitrary axis.

Problem XXV : Show that if the Hamiltonian function (4.2) remains invariant under a linear displacement δr , then the quantity $P \cdot \delta r$, where $P = \sum_n p_n$, is a constant of the motion. Hence, if δr may be chosen arbitrarily, the components of P are individually constants of the motion. That is

$$\dot{P} = 0$$

This is known as the law of conservation of momentum, P being the total momentum of the system.

Suppose that the particle coordinate variables r_n are subject to a certain number of integrable constraints so that the configuration of the assembly is determined by a set of generalized coordinates ρ^i fewer in number than the totality of particle coordinate components. These constraints may be imagined to be imposed by means of infinitely deep but infinitely thin potentials contained in the function V of (4.1, 2), and hence one expects that the Lagrangian function for the constrained system remains essentially that given by (4.1), but expressed in terms of the ρ^i instead of the $\vec{r}_n$. The mathematical verification of this expectation is straightforward. We write

$$r_n = r_n(\rho^i, t) \tag{4.15}$$

$$\dot{r}_n = \frac{\partial r_n}{\partial \rho^i} \dot{\rho}^i + \frac{\partial r_n}{\partial t} \tag{4.16}$$

$$\frac{\partial \dot{r}_n}{\partial \dot{\rho}^i} = \frac{\partial r_n}{\partial \rho^i} \tag{4.17}$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\rho}^i} - \frac{\partial L}{\partial \rho^i} = \sum_n \left\{ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_n} \cdot \frac{\partial r_n}{\partial \rho^i}\right) - \frac{\partial L}{\partial r_n} \cdot \frac{\partial \dot{r}_n}{\partial \rho^i} - \frac{\partial L}{\partial \dot{r}_n} \cdot \frac{\partial r_n}{\partial \rho^i} \right\}$$

$$= \sum_n \left\{ \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{r}_n} - \frac{\partial L}{\partial r_n} \right) \cdot \frac{\partial r_n}{\partial \rho^i} + \frac{\partial L}{\partial \dot{r}_n} \cdot \left(\frac{d}{dt} \frac{\partial r_n}{\partial \rho^i} - \frac{\partial \dot{r}_n}{\partial \rho^i} \right) \right\} \quad (4.18)$$

But

$$\frac{d}{dt} \frac{\partial r_n}{\partial \rho^i} - \frac{\partial \dot{r}_n}{\partial \rho^i} = \frac{\partial^2 r_n}{\partial \rho^\alpha \partial \rho^i} \dot{\rho}^\alpha + \frac{\partial^2 r_n}{\partial t \partial \rho^i} - \frac{\partial^2 r_n}{\partial \rho^i \partial \rho^\alpha} \dot{\rho}^\alpha - \frac{\partial^2 r_n}{\partial \rho^i \partial t} = 0 \quad (4.19)$$

Hence

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\rho}^i} - \frac{\partial L}{\partial \rho^i} = \sum_n \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{r}_n} - \frac{\partial L}{\partial r_n} \right) \cdot \frac{\partial r_n}{\partial \rho^i} = 0. \quad (4.20)$$

This result holds even if the original Lagrangian function has a more general form than that given by (4.1). For example, the potential function V might depend on the velocities $\dot{r}_n$ in addition to the positions r_n, or it might even depend explicitly on the time t.

A particular case of interest is that in which the forces of constraint are such as to maintain all the mutual distances between the particles fixed. The assembly is then said to form a rigid body. In the study of rigid bodies it is useful to make a separation of the total angular momentum into an orbital angular momentum and a spin angular momentum. The discussion which follows will be valid for general assemblies of particles. The specialization to rigid bodies will be made later.

We first introduce the total mass

$$M = \sum_n m_n, \quad (4.21)$$

the center of mass

$$R = \frac{1}{M} \sum_n m_n r_n, \quad (4.22)$$

and the total momentum

$$P = \sum_n p_n \tag{4.23}$$

From (4.3) it is evident that

$$P = M\dot{R}. \tag{4.24}$$

If the Hamiltonian function (4.2) is invariant under arbitrary displacements, then, according to the result of Problem XXV,

$$\dot{P} = 0. \tag{4.25}$$

Introduce now the relative coordinates

$$\rho_n = r_n - R \tag{4.26}$$

and the relative momenta

$$\Pi_n = p_n - \frac{m_n}{M} P = m_n \dot{\rho}_n \tag{4.27}$$

satisfying

$$\sum_n m_n \rho_n = 0 \tag{4.28}$$

and

$$\sum_n \Pi_n = 0 \tag{4.29}$$

Then we may write the total angular momentum in the form

$$\mathcal{J} = \sum_n (R + \rho_n) \times \left(\frac{m_n}{M} P + \Pi_n \right) = L + S \tag{4.30}$$

where

$$L = R \times P \tag{4.31}$$

and

$$S = \sum_n \rho_n \times \Pi_n \tag{4.32}$$

L and S are the orbital and spin angular momenta respectively. Using (4.24) and (4.25) we have

$$\dot{L} = \dot{R} \times P + R \times \dot{P} = \frac{1}{M} P \times P = 0. \qquad (4.33)$$

If the Hamiltonian function (4.2) remains invariant under arbitrary rotations then, combining Eqs. (4.14), (4.30), and (4.33), we obtain

$$\dot{S} = 0. \qquad (4.34)$$

Now, it may be readily verified that

$$(R, \Pi_n) = 0 \,, \quad (\rho_n, P) = 0 \,. \qquad (4.35)$$

Problem XXVI : Verify Eqs. (4.35)

Hence

$$(L \,, S) = 0, \qquad (4.36)$$

and the orbital and spin angular momenta are seen to be dynamically independent, not generating any new constants of the motion.

The above separation of the assembly coordinates into spin and orbital parts is useful even when the laws of conservation of momentum and angular momentum do not apply. This is particularly true of rigid bodies, to which we now turn our attention. We shall consider an especially important example, namely that of a charged rigid body moving under the influence of an impressed static electromagnetic field. If $\mathcal{E}_n$ denotes the charge on the n th particle then we may infer from expression (3.3) of the preceding chapter that the Lagrangian function which describes such a system is given, in the non-relativistic limit, by

$$L = \sum_n \frac{1}{2} m_n r_n^2 + \sum_n \frac{e_n}{c} \dot{r}_n \cdot A(r_n) - \sum_n e_n \varphi(r_n) - V_c \tag{4.37}$$

where V_c is the potential function describing the forces of constraint which hold the body rigid. Since the body is rigid V_c is essentially constant and may be dropped from the theory, provided we remember that the r_n must eventually be expressed in terms of a completely independent and unconstrained set of generalized coordinates ρ^c.

Introducing the coordinates R , r_n defined above, and remembering that

$$\sum_n m_n \dot{\rho}_n = 0, \tag{4.38}$$

we may write (4.37) in the form

$$\begin{aligned} L &= \sum_n \frac{1}{2} m_n (\dot{R} + \dot{\rho}_n)^2 + \sum_n \frac{e_n}{c} (\dot{R} + \dot{\rho}_n)_i \left\{ A_i(R) + \rho_{nj} A_{i,j}(R) \right. \\ &\quad \left. + \frac{1}{2!} \rho_{nj} \rho_{nk} A_{i,jk}(R) + \dots \right\} - \sum_n e_n \left\{ \varphi(R) + \rho_{ni} \varphi_{,i}(R) \right. \\ &\quad \left. + \frac{1}{2!} \rho_{ni} \rho_{nj} \varphi_{,ij}(R) + \dots \right\} \\ &= \frac{1}{2} M \dot{R}^2 + \frac{e}{c} \dot{R} \cdot A(R) - e\varphi(R) \\ &\quad + \sum_n \frac{1}{2} m_n \dot{\rho}_n^2 + \frac{1}{c} \dot{R}_i \eta_j A_{i,j}(R) + \frac{1}{2} \sum_n \frac{e_n}{c} \dot{R}_i \rho_{nj} \rho_{nk} A_{i,jk}(R) \\ &\quad + \frac{1}{c} \dot{\eta} \cdot A(R) + \sum_n \frac{e_n}{c} \dot{\rho}_{ni} \rho_{nj} A_{i,j}(R) - \eta_i \varphi_{,i}(R) \\ &\quad - \frac{1}{2} \sum_n e_n \rho_{ni} \rho_{nj} \varphi_{,ij}(R) + \dots \end{aligned} \tag{4.39}$$

Here we have carried out a Taylor expansion of the scalar and vector potentials about the point R. The quantities e and η are the total charge and the electric dipole moment of the assembly respectively :

$$e = \sum_n e_n , \tag{4.40}$$

$$\eta = \sum_n e_n \rho_n . \tag{4.41}$$

If the variations of the potentials A and φ over the dimensions of the rigid body are small compared to the magnitudes of A and φ themselves, then the body itself may be regarded as small, and we may , with good accuracy, retain only those terms in the Taylor expansions of order up to the second in the ρ_n 's.

If we impose the Lorentz condition (3.32) on the potentials, which in the static case reduces simply to

$$\nabla \cdot A = 0 \tag{4.42}$$

then we have

$$\nabla^2 A = 0 \ , \ \nabla^2 \varphi = 0 \tag{4.43}$$

assuming that the rigid body is not in the vicinity of the impressed charge and current distributions which are producing the impressed electromagnetic field. Under these conditions we may rewrite (4.39) in the form

$$L = \frac{1}{2} M \dot{R}^2 + \frac{e}{c} \dot{R} \cdot A(R) - e\varphi(R)$$
$$+ \sum_n \frac{1}{2} m_n \dot{\rho}_n^2 + \frac{1}{c} \dot{R}_i \eta_j A_{ij}(R) + \frac{1}{c} \dot{\eta} \cdot A(R) - \eta_i \varphi_{,i}(R)$$

$$+\frac{1}{6}\frac{1}{c}\dot{R}_i Q_{jk} A_{i,jk}(R) - \frac{1}{6} Q_{ij}\varphi_{,ij}(R)$$

$$+\frac{1}{c}\sum_n e_n \dot{\rho}_{ni}\rho_{nj} A_{i,j}(R) \qquad (4.44)$$

where Q_{ij} is the electric quadrupole moment of the body, defined in such a way that its trace vanishes :

$$Q_{ij} = \sum_n e_n (3\rho_{ni}\rho_{nj} - \delta_{ij}\rho_n^2) \quad , \quad Q_{ii} = 0. \qquad (4.45)$$

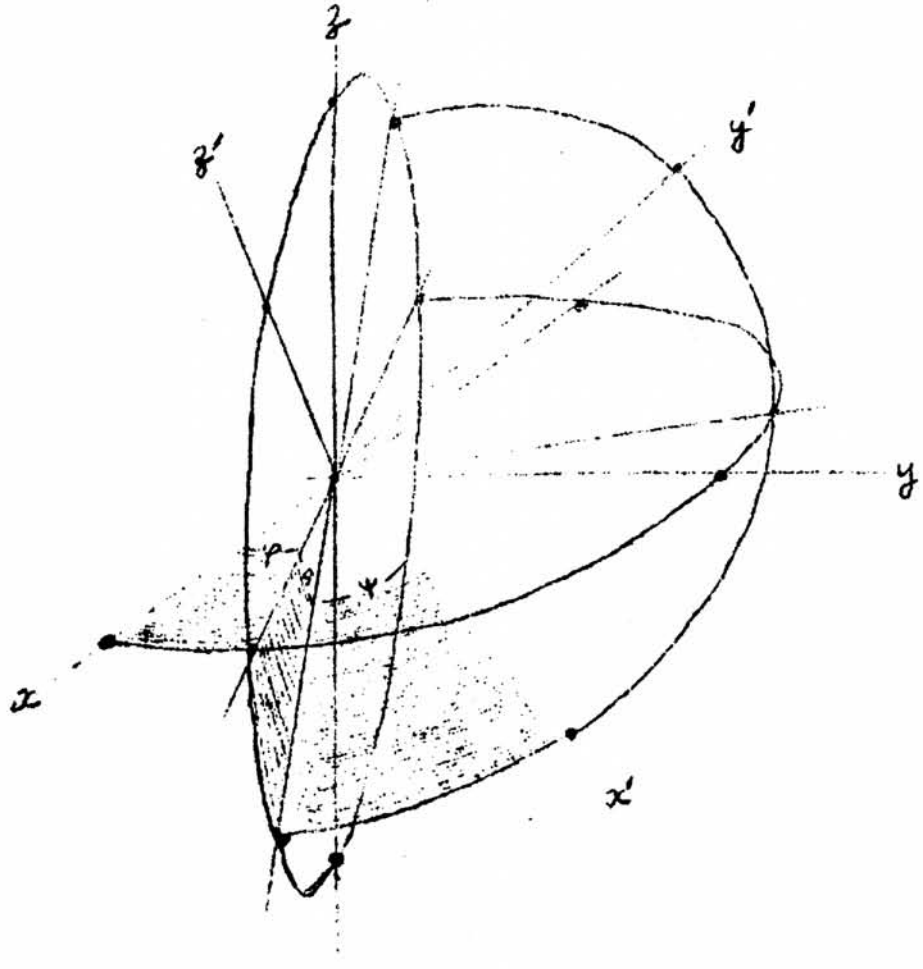
We shall now introduce a special set of six independent generalized coordinates with which to describe the configuration of the rigid body. Three of these will be simply the components of the center of mass R. The other three will be the well known Eulerian angles. Let x', y', z' denote a set of rectangular axes fixed with respect to the body. Let x, y, z, be a set of rectangular axes in space. Then the orientation of the body with respect to its center of mass may be specified by the three angles φ, θ, ψ indicated in the accompanying diagram.

In terms of φ, θ, ψ, the angular velocities of the body about the axes x, y, z are respectively

$$\left.\begin{aligned} \omega_1 &= -\dot{\theta}\sin\varphi + \dot{\psi}\sin\theta\cos\varphi \\ \omega_2 &= \dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi \\ \omega_3 &= \dot{\varphi} + \dot{\psi}\cos\theta \end{aligned}\right\} \qquad (4.46)$$

The angular velocities about the axes x', y', z' are

$$\left.\begin{aligned} \omega'_1 &= -\dot{\varphi}\sin\theta\cos\psi + \dot{\theta}\sin\psi \\ \omega'_2 &= \dot{\varphi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\ \omega'_3 &= \dot{\varphi}\cos\theta + \dot{\psi} \end{aligned}\right\} \qquad (4.47)$$


z
z'
y'
y
x
x'
ρ
φ
ψ

In terms of the angular velocities ω_i, the velocities $\dot{\rho}_m$ are given by

$$\dot{\rho}_m = \omega \times \rho_m \tag{4.48}$$

We shall now specialize to the case in which the rigid body is continuous and has complete spherical symmetry. The radius vector to any point in the body from its center of mass will be denoted by r, and the sums in (4.44) will be replaced by integrals. The charge and mass densities, denoted by ρ_e and ρ_m respectively, will be dependent only on the radial distance r ($= |r|$):

$$\rho_e = \rho_e(r) \quad , \quad \rho_M = \rho_M(r) \tag{4.49}$$

We shall then have

$$\gamma = \int_\infty \rho_e r \, d_3 r = 0 \tag{4.50}$$

and, since,

$$\int_\infty f(r) r_i r_j d_3 r = \frac{1}{3} \delta_{ij} \int_0^\infty f(r) r^2 d_3 r \tag{4.51}$$

$$= \frac{4\pi}{3} \delta_{ij} \int_0^\infty r^4 f(r) dr \tag{4.52}$$

$$Q_{ij} = \int_\infty \rho_e (3 r_i r_j - \delta_{ij} r^2) d_3 r = 0 .$$

That is, the elctric dipole and quadrupole moments vanish. The fourth term of (4.44) becomes

$$\int_\infty \frac{1}{2} \rho_m (\omega \times r)^2 d_3 r = \frac{1}{2} \int_\infty \rho_M \delta_{ijk} \epsilon_{lms} \omega_j r_k \omega_r r_s$$

$$= \frac{1}{3} \omega^2 \int_\infty r^2 \rho_M d_3 r = \frac{1}{2} I \omega^2 \tag{4.53}$$

where I is the moment of inertia of the body

$$I = \frac{2}{3} \int_\infty r^4 \rho_M d_3 r = \frac{8\pi}{3} \int_0^\infty r^4 \rho_M(r) dr \tag{4.54}$$

The last term of (4.44) becomes

$$\frac{1}{c}\int_\infty \rho_e \epsilon_{ijk}\omega_i r_k r_\ell A_{i,\ell}(R)\, d_3 r$$
$$= \frac{1}{3}\,\frac{1}{c}\,\epsilon_{ijk}\omega_j A_{i,k}(R)\int_\infty \rho_e r^2 d_3 r \qquad (4.55)$$

Expressions (4.55) may be rewritten in terms of the magnetic dipole moment of the body, which is defined by

$$\mu = \frac{1}{2c}\int_\infty r \times j\, d_3 r \qquad (4.56)$$

where j is the electric current density. In the present case

$$j = \rho_e \omega \times r \qquad (4.57)$$

Hence

$$\mu_i = \frac{1}{2c}\int_\infty \rho_e \epsilon_{ijk}\epsilon_{krs} r_j \omega_r r_s\, d_3 r$$
$$= \frac{1}{3}\,\frac{1}{c}\,\omega_i \int_\infty \rho_e r^2 d_3 r, \qquad (4.58)$$

Evidently (4.55) may be written in the form

$$\mu \cdot \nabla \times A(R) = \mu \cdot H(R) \qquad (4.59)$$

where H(R) is the magnetic field at the point R.

The magnetic field moment is related to the spin angular momentum. Tha latter is, in the present case, defined by

$$S = \int_\infty \rho_M r \times (\omega \times r)\, d_3 r = I\omega. \qquad (4.60)$$

Comparison of (4.56), (4?57) and (4.60) makes it evident that

$$\mu = \rho S \qquad (4.61)$$

where

$$\rho = \frac{1}{2c}\,\frac{\int_\infty \rho_e r^2 d_3 r}{\int_\infty \rho_M r^2 d_3 r} = \frac{1}{2c}\,\frac{\int_0^\infty \rho_e(r) r^4 dr}{\int_0^\infty \rho_M(r) r^4 dr} \qquad (4.62)$$

ρ is called the gyromagnetic ratio of the body. If the ratio ρ_e / ρ_M

is every where constant then

$$\rho = \frac{e}{2Mc} . \tag{4.63}$$

The Lagrangian function (4.44) may now be written in the form

$$L = \frac{1}{2} M\dot{R}^2 + \frac{e}{c}\dot{R}A - e\varphi + \frac{1}{2} I\omega^2 + \rho I\omega \cdot H \tag{4.64}$$

where it is understood that the electromagnetic quantities are all to be evaluated at the point R. The Hamiltonian function may readily be obtained from (4.64). The momentum conjugate to the center-of-mass coordinate R is now slightly different from that given by (4.24). We have

$$P = \frac{\partial L}{\partial \dot{R}} = M\dot{R} + \frac{e}{c} A. \tag{4.65}$$

The momenta conjugate to the Eulerian angles φ , θ , ψ need not be explicitly calculated in the construction of the Hamiltonian function. Even though the angular velocities ω_i are not integrable (i.e. they are not exact time derivatives of any coordinates), momenta corresponding to them may be introduced :

$$\Pi = \frac{\partial L}{\partial \omega} = I\omega + \rho IH = S + \rho IH \tag{4.66}$$

The momenta Π_i are related to the momenta p_φ, p_θ, p_ψ conjugate to the Eulerian angles, inversely as the angular velocities ω_i are related to the velocities $\dot{\varphi}$, $\dot{\theta}$, $\dot{\psi}$. These relations are conveniently expressed in matrix form. Referring to (4.46), we have

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 0 & -\sin\varphi & \sin\theta\cos\varphi \\ 0 & \cos\varphi & \sin\theta\sin\varphi \\ 1 & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} \dot{\varphi} \\ \dot{\theta} \\ \dot{\psi} \end{pmatrix} \tag{4.67}$$

$$(\Pi_1\ \Pi_2\ \Pi_3) = (p_\varphi\ p_\theta\ p_\psi) \begin{pmatrix} \cos\varphi\cot\theta & -\sin\varphi\cot\theta & 1 \\ -\sin\varphi & \cos\varphi & 0 \\ \cos\varphi\csc\theta & \sin\varphi\csc\theta & 0 \end{pmatrix} \tag{4.68}$$

The matrix appearing in (4.68) is the inverse of that appearing in (4.67).

We may now calculate the Hamiltonian function as follows :

$$H = P\cdot\dot{R} + \Pi\cdot\omega - L$$

$$= \left(P - \frac{e}{c}A\right)\cdot\dot{R} - \frac{1}{2}MR^2 + (\Pi - \rho[I]H)\cdot\omega - \frac{1}{2}I\omega^2 + e\varphi$$

$$= \frac{1}{2M}\left(P - \frac{e}{c}A\right)^2 + \frac{1}{2I}(\Pi - \rho[I]H)^2 + e\varphi \tag{4.69}$$

In using this Hamiltonian function to obtain the equations of motion, one must remember that since the Π_i correspond to non-integrable velocities they do not have vanishing Poisson brackets with one another as they would have if they were conjugate to real coordinates. The Poisson bracket relations which they satisfy may be inferred from a consideration of the case in which the impressed electromagnetic field is not present. The Lagrangian function then reduces to

$$L = 1/2\ M\ \dot{R}^2 \ +\ 1/2\ I\ \omega^2 \tag{4.70}$$

and the momenta corresponding to ω becomes simply the spin angular momentum

$$\Pi = I\omega = S\ . \tag{4.71}$$

The orbital angular momentum is given by expression (4.31) and satisfies, by itself, Poisson bracket relations analogous to these of (4.11) satisfied the total angular momentum, namely

$$(L_i, L_j) = \varepsilon_{ijk} L_k . \tag{4.72}$$

Since the Hamiltonian function is now invariant under arbitrary rotations, equations (4.11), (4.36), and (4.72) together imply

$$(S_i, S_j) = \varepsilon_{ijk} S_k . \tag{4.73}$$

This may be rewritten in the form

$$(\Pi_i, \Pi_j) = \varepsilon_{ijk} \Pi_k . \tag{4.74}$$

Now the Poisson bracket relations satisfied by the Π_i depend only on the form of the linear relation (4.68) connecting them with the Eulerian momenta $p_\varphi, p_\theta, p_\psi$. Hence relations (4.74) will be satisfied even when an electromagnetic field is present.

Problem XXVII : Verify (4.74) directly from (4.68)

From (4.74) we obtain $(\Pi_i, \Pi^2) = 2\Pi_j \varepsilon_{ijk} \Pi_k = 0$.

Moreover, the dynamical equations give

$$\dot{\Pi} = (\Pi, H) = \left(\Pi, \frac{1}{2I}\Pi^2 - \rho \Pi \cdot H + \frac{1}{2}\rho^2 I H^2\right) = -\rho(\Pi, \Pi, H)$$

$$= \rho \Pi \times H \tag{4.76}$$

so that

$$\frac{d}{dt}\Pi^2 = 2\dot{\Pi}\cdot\Pi = 0. \tag{4.77}$$

Thus the term $\frac{1}{2I}\Pi^2$ occurring in the Hamiltonian, having no effect

on the dynamical equations since it has vanishing Poisson bracket with R, P, and Π, and being a constant of the motion by (4.77) may be dropped from the Hamiltonian, leaving

$$H = \frac{1}{2M}\left(P - \frac{e}{c}A\right)^2 - \rho\,\Pi\cdot H + \frac{1}{2}\rho^2 I H^2 + e\varphi \qquad (4.78)$$

This procedure corresponds to an energy renormalization in which the spin rotational energy of the body is substracted out.

We shall now pass to a conceptual limit in which the body becomes infinitesimally small. In this case, if the mass remains finite it is evident that the moment of inertia I must tend to zero. In order to maintain a finite spin angular momentum the angular velocity ω must therefore tend to infinity :

$$I \to 0 \text{ and } \omega \to \infty \text{ in such a way that } I\omega = S. \qquad (4.79)$$

Evidently under these conditions

$$\Pi \to I\omega = S \qquad (4.80)$$

$$\frac{1}{2I}\Pi^2 \to \frac{1}{2} I\omega^2 \to \infty \qquad (4.81)$$

and the energy renormalization mentioned above becomes an infinite one . The Hamiltonian function becomes

$$H = \frac{1}{2M}\left(P - \frac{e}{c}A\right)^2 + e\varphi - \rho S\cdot H \qquad (4.82)$$

and the Poisson bracket relations for S reduce to (4.73).

In the case of an infinitesimally small body the higher order terms in the Taylor expansion of (4.39) become infinitesimal and the Hamiltonian function (4.82) becomes exact. Reviewing what we have done in this chapter,, we see that we have started with the concept of charged mass points, and out of this constructed the notion of finite sized, rigid material bodies. Then, by extrapolating

to the case of infinitesimally small bodies, we have come full circle back again to the notion of point particles. But now our particles are endowed with one more property than they had before, namely spin angular momentum. The existence of this new property is reflected in the presence of the last term in the Hamiltonian function (4.82). This extra term, together with the Poisson bracket relations (4.73), suffice completely to describe the spin angular momentum of the particle.

Strictly speaking there is, in principle, one more property possessed by any spinning spherical body, namely the orientation of the body axes x', y', z'. If the spin angular momentum, and hence the angular velocities ω_i, are given, then the orientation of the body axes will be determined by giving simultaneously the angular velocities ω_i' relative to these axes.

---

* Problem XXVIII : Defining $S' = I\omega'$, show that

$$(S_i', S_j') = -\epsilon_{ijk} S_k'$$

$$S'^2 = S^2, \text{ and } (S, S') = 0.$$

However, if all points on the spherical surface of the body (or all points on any concentric spherical shell) are completely indistinguishable, then to speak of the orientation of the body is conceptually meaningless. Or at any rate the orientation of the body is completely unobservable. The orientation certainly cannot be observed through interactions described by the Hamiltonian function (4.82). Some other interacting system would have to be introduced to make the observation.

As far as the Hamiltonian function (4.82) is concerned, therefore, the concept of the orientation of the body axes x', y',

$3'$, may be dropped from the theory. Instead of six internal observables, e.g. the three internal coordinates (Eulerian say) and their conjugate momenta, there remain only three, namely the components of the spin angular momentum. An understanding of this point will prepare the way for the later acceptance, in the quantum theory, of the existence of dynamical observables for which there are no canonically conjugated quantities.

5.- THE CORRESPONDANCE PRINCIPLE AND THE OLD QUANTUM THEORY.

A great wealth of experimental evidence points to the fact that matter and radiation do not obey classical laws in the atomic domain. In particular, one of the first facts inferred historically was that energy is not transferred from one to another of two interacting systems in a continuous manner, but by a series of discrete jumps of well defined amount. Particularly, illuminating is the case of a multiply periodic, charged, bound system (e.g. an atom, molecule or crystal) interacting either with electromagnetic radiation or with moving, charged particles. If a beam of moving charged particles, of definite kinetic energy E, is directed at the multiply periodic system it is found that the particles are either scattered elastically, in which case there is no transfer of energy (assuming the system to be much heavier than the particles), or else they are scattered inelastically with very definite discrete drops, ΔE_i , in energy. These energy changes ΔE_i can be very well recorded, and appear to be constants characteristic of the particular multiply periodic system in question. Presumably the discrete energy "quanta", ΔE_i , are transferred from the moving particles to the system.

Again, if the multiply periodic system is once excited, e.g. by collisions with moving particles, as above, it will radiate owing to the motion of its component parts. It is found that the emitted electromagnetic radiation does not possess a continuum of frequencies, but rather a spectrum of well defined discrete angular frequencies, ω_i . These angular frequencies, moreover, are observed to stand in a very definite relationship to the energy changes mentioned above, namely

$$\Delta E_i = \hbar \omega_i \qquad (5.1)$$

where $\hbar$ is a universal constant which was first introduced by M. Planck.

One would be inclined to suggest that the radiation process described above is one in which the energy quanta ΔE_i, having been stored in the multiply periodic system, are once again released, this time to the electormagnetic field, the energy transfer process being as before a discrete process. Not only is this in fact the case, but experiments on the photoelectric effet show that radiation angular frequencies of amount ω_i are inseparably connected with energy units ΔE_i given by (5.1). For example; monochromatic light of angular frequency ω_i falling on a photoelectric metallic surface can impart energy to the electrons in the metal only in units of $\hbar\omega_i$.

Now, monochromatic radiation can be regarded as a periodic system. Therefore it is a logical question to ask whether other periodic systems, in particular the multiply periodic system considered above, store energy units in a manner analogous to that of radiation. More precisely, is there a relationship between the various discrete amounts of energy which a multiply periodic system can store and the angular frequencies of the system ? We shall see presently that there is a relationship between these quantities, but that it is, in general, not quite so simple as that expressend by (5.1).

We first observe that not all classical concepts need be renounced in attempting to deal with the phenomena described above. For example, the concept of conservation of energy is quite logically retained. Very close account can be kept of the energy as it is transferred from one system to another -from the moving particles, to the multiply periodic system, to the electromagnetic radiation and finally to the electrons in the photoelectric cell. The only unusual feature is that the energy is now compelled to be transferred in definite discrete units.

Secondly, it is evident, from the observed classical behavior of macroscopic matter, that the behavior of any multiply periodic system must become more and more nearly classical as its energy content increases to magnitudes much larger than $\hbar\omega^i$ where ω^i

ω^i is a typical angular frequency of the system, because the individual energy jumps will then become negligible compared to the energies with which one is dealing, and the dynamical processes will become effectively continuous and classical. This principle, which was first enunciated by N. Bohr, is known as the Correspondence Principle. It is through the use of this principle that we shall be guided to the relationship between the frequencies of a multiply periodic system and its quantum energy levels.

Now, the angular frequencies of a multiply periodic system are given by

$$\omega^i = \frac{\partial E}{\partial J_i} \tag{5.2}$$

(see(1.161)), where E is the energy of the system and the J_i are the non-degenerate action variables. According to the classical theory the electromagnetic field radiated by the system will in general involve all the angular frequencies ω^i together with their higher harmonics. By the Correspondance Principle this will also be true in the quantum theory in the classical limit of high energies. Now, a very plausible explantion of the fact that the multiply periodic system can store energy only in characteristic units ΔE_i is to assume that the system can exist in only certain discrete energy levels E_m. The units ΔE_i would then represent simply the differences between the various energy levels. Thus, for a transition from the m th energy level to the n th energy level, the energy quantum E_{nm} involved would be given by

$$\Delta E_{nm} = E_n - E_m \tag{5.3}$$

E_{nm} can be either positive or negative according to whether the quantum is absorbed or emitted. In the case of emission the transition is from the higher level, say E_m, to the lower level, E_n, and by (5.1) the angular frequency of the emitted radiation is

$$\omega_{nm} = \frac{\Delta E_{nm}}{\hbar} = \frac{E_n - E_m}{\hbar} \qquad (5.3)$$

The fact that ω_{nm} is negative for emission is unimportant as only the magnitude of ω_{nm} is of significance.

In the limit of high energies this must be very closely equal to one of the harmonics of the classical frequencies .

Let us assume that the system is non-degenerate, or, rather, that it is at most accidentally degenerate. If the system is degenerate it can always be made non-degenerate by introducing small perturbations which remove the degeneracy but which do not change the physical characteristics very much, and hence we lose no generality by this assumption. Then the energy must be specified by a complete set of action variables J_i , one for each degree of freedom of the system. Corresponding to the n th energy level there will be a set of values J_{ni} for the action variables. More properly, therefore, we should denote the energy levels by $E_{n_1 \ldots n_N}$, where N is the number of degrees of freedom and

$$E_{n_1 \ldots n_N} = E(J_{n_1} \ldots J_{n_N}). \qquad (5.4)$$

However, it will be convenient to continue to use the subscripts n, m, etc. as shorthand for $n_1 \ldots n_N$, $m_1 \ldots m_N$, etc. just as in chapter 2 we used the subscript τ as shorthand for $\tau_1 \ldots \tau_N$. The indices n_i have thus far been introduced purely for labeling purposes and may conveniently be taken to have the integral values 0, 1, 2, 3,.... and possibly also -1, -2, -3,... The n_i will be known as <u>quantum numbers</u>.

We may now write

$$\Delta E_{nm} = E(J_n) - E(J_m)$$
$$= \frac{\partial E}{\partial J_i}(J_m)\,\Delta J_{nmi} + \frac{1}{2!}\frac{\partial^2 E}{\partial J_i \partial J_j}(J_m)\,\Delta J_{nmi}\,\Delta J_{nmj} + \ldots$$

$$= \omega^i(J_m)\,\Delta J_{nmi} + \frac{1}{2}\frac{\partial \omega^i}{\partial J_j}(J_m)\,\Delta J_{nmi}\,\Delta J_{nmj} + \ldots \tag{5.5}$$

where

$$\Delta J_{nmi} = J_{ni} - J_{mi}. \tag{5.6}$$

Since E is in general a monotonic increasing function of the J's, the classical limit is given by high J's. For fixed Δ J's, the second term of the expansion (5.5) in general becomes negligible in comparison with the first term as the J's become increasingly large. Hence in the classical limit we may write, combining (5.3) and (5.5),

$$\hbar\omega_{nm} = \Delta E_{nm} \approx \omega^i \Delta J_{nmi} \tag{5.7}$$

where ω^i may be evaluated either at Jm or Jn.

Problem XXIX : Show that for the case of a particle in a box (see Problem XX),

$$\lim_{J\to\infty} \frac{\partial \omega}{\partial J}\frac{\Delta J}{\omega} = 0$$

But since ω_{nm} must be nearly a harmonic of the system, we have

$$\omega_{nm} \approx \tau_i \omega^i \tag{5.8}$$

where the τ_i are certain intergers. Equations (5.7) and (5.8) are evidently compatible if

$$\Delta J_{nmi} = \hbar\tau_i. \tag{5.9}$$

The first historic attempt to construct a systematic quantum theory consisted in postulating that Eq. (5.9) was not only valid in the classical domain but could also be extrapoled right down into the quantum domain. Evidently, integral relations of the form (5.9) can exist between every pair of quantum energy levels only if

$$\mathcal{J}_{ni} = (n_i + \gamma_i)h \qquad (5.10)$$

where the n_i are the integral quantum numbers and the γ_i are certain constants. Equations (5.10) are known as the Bohr-Sommerfeld quantum conditions.

The constants γ_i are as yet undetermined by the theory. Actually, the original Bohr-Sommerfeld hypothesis set these constants equal to zero. We may anticipate later developments somewhat, however, by stating that experiments show that these constants are frequently not zero, but are nevertheless determined by certain simple general rules. Generally speaking,

A) $\gamma_i = 0$ if the associated action variable corresponds to a motion of rotation or of libration in a box with vertical potential walls. In the latter case n_i is restricted to the values 1, 2, 3,... while in the former case n_i may take on all positive and negative integral values.

B) $\gamma_i = 1/2$ if the associated action variable corresponds to a motion of libration in a smoothly varying potential. In this case n_i is restricted to the values 0, 1, 2, 3,...

Other values for these constants are occasionally encountered in cases which fall somewhat between A) and B). However, in such cases reference is best made to the exact quantum mechanics which will be developed in succeeding chapters.

We shall now apply the Bohr-Sommerfeld conditions to calculate the quantum energy levels for some of the simple systems considered in preceding chapters. The prototype of a periodic system is the harmonic oscillator. From equation (1.185) we have for the energy of a 1-dimensional harmonic oscillator

$$E = J\omega. \tag{5.11}$$

Since the potential function, namely $1/2\, m\omega^2\rho^2$, is smoothly varying, we have case B) above, and hence the energy levels are given by

$$E_n = (n + 1/2)\, \hbar\omega. \tag{5.12}$$

It will immediately be observed that if transitions are assumed to take place only between adjacent levels, then the energy quanta are all given by

$$\Delta E = \hbar\omega. \tag{5.13}$$

This is the relation we have already seen to exist in the case of radiation.

In the case of the plane rotator, which falls under A) above, we have, using Eq. (1.206),

$$E_n = \frac{1}{2I} n^2 \hbar^2 \tag{5.14}$$

Using (1.207) we find that the frequencies corresponding to these energy levels are all harmonics of a basic frequency $\frac{\hbar}{I}$:

$$\omega_n = \frac{n\hbar}{I} = n\omega_1 \quad \text{where } \omega_1 = \frac{\hbar}{I}. \tag{5.15}$$

The frequencies corresponding to actual transitions, however, are not exact harmonics. For a radiative transition from the $(n+1)$ th level to the n th level we have

$$\omega_{n+r,n} = \frac{E_{n+r} - E_n}{\hbar} = \frac{\hbar}{2I}\left[(n+r)^2 - n^2\right]$$

$$= \frac{\hbar}{I} r\left(n + \frac{r}{2}\right) = r\,\frac{\omega_{n+r} + \omega_n}{2} \qquad (5.16)$$

Thus the emitted frequency is the r th harmonic of a frequency which lies exactly half way between the frequencies corresponding to the two levels. This situation is typical of all quantized periodic systems, although the emitted frequency will in general not be expressible as a simple arithmetic average. As n becomes very large it is readily seen that the emitted frequencies do approach harmonics of the fundamental frequency ω_1.

Problem XXX : Show for the one-dimensional particle in a box (see Problem XX) that the energy levels are given by

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{\ell}\right)^2 \qquad n = 1, 2, 3, \ldots$$

A very important example of an easily quantizable system is that of a non-relativistic particle moving in a central field of force. We shall assume that the inherent degeneracy of such a system has been removed by small perturbations. Then, referring, to Eqs. (1.228, 229, 230), we have

$$J_\varphi = n_\varphi \hbar \qquad (5.17)$$

$$J_\theta = \left(n_\theta + \tfrac{1}{2}\right)\hbar \qquad (5.18)$$

$$J_r = \left(n_r + \tfrac{1}{2}\right)\hbar. \qquad (5.19)$$

J_φ corresponds to a motion of rotation and hence comes under case A) above. Moreover, since J_φ can be either positive or negative, depending on the direction of the rotation, n_φ can take on all positive or negative integral values. J_θ and J_r, on the other hand, correspond to motions of libration in smoothly varying potentials and hence come under case B). Moreover, the inequalities in (1.229) and (1.230) indicate that

$$n_\theta \geq 0 \quad \text{and} \quad n_r \geq 0 . \tag{5.20}$$

If we transform to the action variables J_n, J_ℓ, J_m (see Eqs. (1.233, 234, 235) and also (1.240, 241), we have

$$J_n = n\hbar \tag{5.21}$$

$$L = J_\ell = (\ell + 1/2)\hbar \tag{5.22}$$

$$L_z = J_m = m\hbar \tag{5.23}$$

where

$$n = |n_\varphi| + n_\theta + n_r + 1, \tag{5.24}$$

$$\ell = |n_\varphi| + n_\theta , \tag{5.25}$$

$$m = n_\varphi , \tag{5.26}$$

The inequalities (1.243), taken together with (5.20), impose the following restrictions on these quantum numbers :

$$n \geq 1 \tag{5.27}$$

$$\ell = 0, 1, 2 \ldots n-1 \tag{5.28}$$

$$m = -\ell, -\ell+1 \ldots -2, -1, 0, 1, 2, \ldots, \ell-1, \ell \quad (5.29)$$

In the case of Kepler motion we find, on referring to (1 (1.260), that the energy levels are given by

$$E_n = -\frac{mZ^2e^4}{2\hbar^2}\frac{1}{n^2} \quad (5.30)$$

The angular frequencies of emitted radiation are therefore

$$\omega_{n'n} = \frac{E_{n'} - E_n}{\hbar} = \frac{mZ^2e^4}{2\hbar^3}\left(\frac{1}{n^2} - \frac{1}{n'^2}\right). \quad (5.31)$$

Spectroscopists usually reexpress this equation in terms of the wavelengths $\lambda_{n'n}$ of the emitted radiation

$$\lambda_{n'n}^{-1} = \frac{\omega_{n'n}}{2\pi c} = Z^2R\left(\frac{1}{n^2} - \frac{1}{n'^2}\right) \quad (5.32)$$

where $$R = \frac{me^4}{4\pi\hbar^3 c} \quad (5.33)$$

R is known as the Rydberg constant. For fixed values of n , the various associated wavelengths (varying n') give rise to the well known spectral series :

n = 1	Lyman series
n = 2	Balmer series
n = 3	Paschen series
n = 4	Bracket series

We saw previously that Kepler motion is 2-fold degenerate. The term "degenerate" is also applied in quantum theory to the individual energy levels E_n. Each set of quantum numbers is said to define one quantum state. It may happen that a number of different states corresponds to the same energy value. If r is the number of these states then the energy level is said to be r -fold degenerate. In the

Kepler case, for each energy value (i.e. for each n-value), the quantum number ℓ can take on n different values (see (5.28)), and for each ℓ the quantum number m can take on $2\ell + 1$ different values (see (5.29)). Hence the degeneracy of the level E_n is given by

$$\nu_n = \sum_{\ell=0}^{n-1} (2\ell+1) = 1+3+5+\ldots+2n-1 = n^2 \qquad (5.34)$$

In the case of relativistic Kepler motion, the results of problem XXI show that the energy levels are given by

$$E_{n,\ell} = \frac{m_0 c^2}{\sqrt{1 + \frac{Z^2\alpha^2}{n - (\ell+1/2) + \sqrt{(\ell+1/2)^2 - Z^2\alpha^2}}}}$$

$$\approx mc^2 - \frac{mZ^2e^4}{2\hbar^2}\frac{1}{n^2}\left[1 + \frac{\alpha^2}{n^2}\left(\frac{n}{\ell+1/2} - \frac{3}{4}\right)\right] \qquad (5.35)$$

where

$$\alpha = \frac{e^2}{\hbar c} \qquad (5.36)$$

When e refers to the charge on an electron then α is known as the fine structure constant and is experimentally observed to have the approximate value 1/137. From Eq. (5.35) it will be observed that the bound state energy values can become complex for $Z\alpha > 1/4$, or $Z > \frac{1}{2\alpha} \approx 68.5$. The physical interpretation of this situation is quite complicated and will not be discussed here.

Equation (5.35) was first obtained by Sommerfeld in an attempt to explain the fine structure of spectral lines. Actually, Sommerfeld's equation had simply ℓ everywhere (5.35) contains $\ell + 1/2$. Hence it was incorrectly derived. However, it agreed with experiment.

That is, Sommerfeld obtained a correct equation by incorrect reasoning. The fact that his equation was actually correct can only be explained by including the effects of electron spin, as we shall see later.

In both the relativistic and non-relativistic central force motions the action variables are given by Eqs. (5.21-23), and the restrictions on the quantum numbers are given by Eqs. (5.24, 26). Of some interest is the case in which $\ell = 0$. Then, according to (5.29), we must have $m = 0$, and (5.23) tells us that the z-component of the angular momentum must vanish. However, the direction of the z-axis is entirely arbitrary, and since we shall see later that expressions (5.22, 23) define the actual physically observable values of the quantities in question (which, like the energy, turn out to be discretely quantized), we could logically say that the total angular momentum must vanish, since its component in any direction vanishes. However, (5.22) tells us that $L = \hbar/2$, and hence $L^2 = 1/4\ \hbar^2$. It will turn out later that both L_z and L^2 are simultaneously observable, and for $\ell = 0$ it will turn out that the measurable values of both L_z and L^2 are zero. Hence in order to remove the residual value $1/4\ \hbar^2$ from the square of the total angular momentum, we should define

$$L^2 = p_\phi^2 - \frac{1}{4}\hbar^2 = \ell(\ell+1)\hbar^2 \qquad (5.37)$$

The residual angular momentum $1/2\ \hbar$ is somewhat analogous to the residual or "zero point" energy, $\frac{1}{2}\omega\hbar$, of a harmonic oscillator in its lowest state (set $n = 0$ in (5.12)).

It is of interest to determine the approximate sizes of the quantized Kepler orbits. From Eq. (1.252) it may be seen that the radial distance r of the moving particle oscillates between two values which are the roots of the equation

$$2mE + \frac{2mZe^2}{r} - \frac{p_\phi^2}{r^2} = 0 \qquad (5.38)$$

The mean value of these roots is given by

$$\frac{1}{2}(r_{max}+r_{min}) = -\frac{1}{2}\frac{Ze^2}{E} = \frac{n^2}{Z} r_B \tag{5.39}$$

where

$$r_B = \frac{\hbar^2}{me^2}. \tag{5.40}$$

If e and m are respectively the electronic charge and mass then r_B is called the Bohr radius.

Thus far we have used the Correspondance Principle only to determine the quantum energy levels of a multiply periodic system. We shall next use it to investigate the radiation of electromagnetic energy from such systems. By far the bulk of the radiation from atomic systems arises from the time variation of their electric dipole moments. To very good approximation the radiation due to magnetic dipole, electric quadrupole, or higher multipole moments may be neglected. Using Eq. (3.177), therefore, we have for the rate of radiation of electromagnetic energy from a multiply periodic atomic system

$$\frac{dE^{rad}}{dt} = \frac{2}{3c^3}\left(\frac{d^2\eta}{dt^2}\right)^2 \tag{5.41}$$

where η is the electric dipole moment. Strictly speaking, Eq. (5.41) is the non-relativistic limit of the exact equation (3.176). This limit is valid here, however, since the velocities of the component parts of the system are assumed to be small compared to the speed of light.

In the classical theory the dipole moment can be expressed in terms of the canonical coordinates ρ^ν, and therefore is a multiply periodic function of the angle variables ω^ν. In general, therefore, there will exist an expression for η of the form

$$\eta = \sum_\tau \eta_\tau(J)\, e^{i\tau\omega} = \sum_\tau \eta_\tau e^{i\tau\omega t}, \tag{5.42}$$

and the mean rate of energy radiation may be expressed as

$$\overline{\frac{dE^{rad}}{dt}} = \frac{4}{3c^3} \sum_{\tau\omega>0} (\tau\omega)^4 |\eta_\tau|^2. \qquad (5.43)$$

We see that, classically, the angular frequencies of the emitted radiation are the harmonics of the fundamental angular frequencies ω^i. In the quantum theory, however, if the system is initially in the m th energy state, it can radiate energy only by going to lower energy states m, the angular frequencies of the associated radiation being given by (5.3). One would therefore expect expression (5.43) to be replaced by an expression of the form

$$\overline{\frac{dE^{rad}}{dt}} = \frac{4}{3c^3} \sum_{E_n < E_m} (\omega_{nm})^4 |\eta_{nm}|^2 \qquad (5.44)$$

where η_{nm} is some average between the values of the Fourier amplitude η_{m-n} at the two energies E_m and E_n :

$$\eta_{nm} \approx \frac{1}{2} \left[\eta_{m-n}(J_m) + \eta_{m-n}(J_n) \right]. \qquad (5.45)$$

Expression (5.44) satisfies the Correspondence Principle since it passes over into (5.43) in the classical limit of high energies E_m and relatively small jumps ΔE_{nm}.

We have not yet developed a theory of how to calculate the true values of the quantities η_{nm}. That will come later. However, we <u>can</u> give definite values to them in certain instances. For example, if for a certain value of τ , η_τ is constant for all values of the action variables J_i , then we can say that $\eta_{m-\tau\, m}$ is equal to that constant for all m. In particular if η_τ is identically equal to zero then $\eta_{m-\tau\, m}$ must vanish for all m. This means that the radiative transitions $E_m \longrightarrow E_{m-\tau}$ are forbidden. Or rather, they cannot take place by means of dipole radiation. They may not be forbidden by higher electric or magnetic multipole radiation, but their corresponding spectral lines will in general be very weak.

Also, Eq. (5.45) suggests, when combined with the reality condition on η , that the following relation between the η_{nm}'s may

be a very good guess :

$$\eta_{rm} = (\eta_{mr})^* . \tag{5.46}$$

Let us now apply the foregoing considerations to a few simple cases.

I - The harmonic oscillator :

The electric dipole moment is simply $\eta = e\varphi$, so that using (2.68), we have

$$\eta = e\sqrt{\frac{\hbar}{2m\omega}}\left(e^{i\omega} + e^{-i\omega}\right) \tag{5.46}$$

and hence

$$\eta_1 = \eta_{-1} = e\sqrt{\frac{\hbar}{2m\omega}} \quad , \quad \eta_\tau = 0 \quad \text{for} \quad \tau \neq \pm 1 . \tag{5.47}$$

In quantum form this becomes

$$\eta_{n\,n-1} = \eta_{n-1\,n} \approx \frac{1}{2} e\left[\sqrt{\frac{j_n}{2m\omega}} + \sqrt{\frac{j_{n-1}}{2m\omega}}\right]$$

$$= e\left[\sqrt{\frac{(n+\frac{1}{2})\hbar}{2m\omega}} + \sqrt{\frac{(n+\frac{1}{2})\hbar}{2m\omega}}\right]$$

$$\approx e\sqrt{\frac{\hbar}{2m\omega}}\sqrt{n} \tag{5.48}$$

We shall see later that this last form is exact. It can be made plausible by the observation that there is no -1st energy so that $\eta_{-1\,0}$ should vanish. That is, no transitions below the ground state can take place. We also have

$$\eta_{n,n-\tau} = 0 \quad \text{for } \tau \neq \pm 1. \tag{5.49}$$

Hence, in a radiative transition, the quantum number n of a harmonic oscillator can only change by $\mp$ 1. This gives us the "selection rule":

$$\Delta n = \pm 1. \tag{5.50}$$

The rate of radiation of energy from the n th level is evidently

$$\frac{d\bar{E}_{rad}}{dt} = \frac{2e^2}{3mc^3}\omega^3 \hbar n = \frac{2e^2}{3mc^3}\omega^2 \frac{E_n + E_{n-1}}{2}. \tag{5.51}$$

This is to be compared with the classical result

$$\frac{d\bar{E}_{rad}}{dt} = \frac{2e^2}{3mc^3}\omega^2 E \tag{5?52}$$

where E is the energy of the oscillator.

Problem XXXI : Prove Eq. (5.52).

II) The plane rotator :

Assume $I = mr^2$, $\eta = cr$, $r = (x, y)$

Then, from (1.209) we have

$$x = r\cos\varphi = r\cos\omega = \frac{r}{2}\left(e^{i\omega} + e^{-i\omega}\right) \tag{5.53}$$

$$y = r\sin\varphi = r\sin\omega = \frac{r}{2i}\left(e^{i\omega} - e^{-i\omega}\right), \tag{5.54}$$

yielding

$$|\eta_1|^2 = |\eta_{-1}|^2 = \frac{e^2\hbar^2}{2} = \frac{e^2}{2m} I = \frac{e^2}{2m\omega} \mathcal{J} \tag{5.55}$$

$$|\eta_\tau|^2 = 0 \quad \text{for } \tau \neq \pm 1 \tag{5.56}$$

where use has been made of (1.205). In the quantum theory, therefore, we have

$$|\eta_{n\,n+\tau}|^2 = 0 \quad \text{for } \tau \neq \pm 1, \tag{5.57}$$

$$|\eta_{n\,n-1}|^2 = |\eta_{n-1\,n}|^2 = \frac{e^2\hbar^2}{2} = \frac{e^2}{2m} I \tag{5.58}$$

We evidently have the selection rule

$$\Delta n = \pm 1 \tag{5.59}$$

just as in the case of the harmonic oscillator. The rate of radiation of energy from the n th level is

$$\begin{aligned} \overline{\frac{dE}{dt}}^{rad} &= \frac{4}{3c^3} \omega_{n-1\,n}^4 |\eta_{n-1\,n}|^2 \\ &= \frac{4}{3c^3} \frac{\hbar^4}{I^4} \left(n - \frac{1}{2}\right)^4 \frac{e^2}{2m} I \\ &= \frac{8}{3c^3} \frac{e^2}{mI} \left(\frac{\sqrt{E_n} + \sqrt{E_{n-1}}}{2} \right)^4 \end{aligned} \tag{5.60}$$

where use has been made of (5.14) and (5.16). Eq. (5.60) is to be compared with the classical result

$$\overline{\frac{dE}{dt}}^{rad} = \frac{8}{3c^3} \frac{e^2}{mI} E^2 \tag{5.61}$$

where E is the energy of the rotator.

Problem XXXII : Prove Eq. (5.61).

Problem XXXIII : Using the results of problems XX and XXX, show that η_{nm} is given unambiguously by

$$\eta_{nm} = \eta_{mn} = \frac{e\ell}{(n-m)^2\pi^2}\left[1-(-1)^{n-m}\right].$$

Show that the rate of energy radiation into the various possible frequencies is given quantum mechanically by

$$\left(\frac{dE^{rad}}{dt}\right)_{nm} = \frac{16}{3c^3}\left(\frac{\sqrt{E_n}+\sqrt{E_m}}{2}\right)^4 \frac{e^2}{m^2\ell^2}\left[1-(-1)^{n-m}\right]$$

and classically by

$$\left(\frac{dE^{rad}}{dt}\right)_{\tau} = \frac{16}{3c^3}E^2\frac{e^2}{m^2\ell^2}\left[1-(-1)^{\tau}\right].$$

The following selection rule is evident

$$\Delta n = \pm 1, \pm 3, \pm 5, \ldots$$

III) The particle moving in a central force field.

Using the results contained in equations (1.222) to (1.235) we may write

$$W_\varphi = J_m \varphi, \tag{5.62}$$

$$W_\theta = \int \sqrt{J_\ell^2 - \frac{J_m^2}{\sin^2\theta}}\, d\theta, \tag{5.63}$$

$$W_r = \int \sqrt{2m\left[E - V(r)\right] - \frac{J_\ell^2}{r^2}}\, dr, \tag{5.64}$$

where it must be remembered that the energy E is a function of J_ℓ and $J_n - J_\ell$. We thus obtain, for the angle variables canonically conjugate to the action variables J_m, J_ℓ , J_r ,

$$\omega_m = \frac{\partial W}{\partial J_m} = \frac{\partial W_\varphi}{\partial J_m} + \frac{\partial W_\theta}{\partial J_m}$$

$$= \varphi - \int \frac{J_m \, d\theta}{\sin^2\theta \sqrt{J_c^2 - \frac{J_m^2}{\sin^2\theta}}}$$

$$= \varphi - \sin^{-1}\left(\frac{J_m}{\sqrt{J_c^2 - J_m^2}} \cot\theta \right), \qquad (5.65)$$

$$\omega_c = \frac{\partial W}{\partial J_c} = \frac{\partial W_\theta}{\partial J_c} + \frac{\partial W_r}{\partial J_c}$$

$$= \int \frac{J_c \, d\theta}{\sqrt{J_c^2 - J_m^2 \csc^2\theta}} + \frac{\partial W_r}{\partial J_c}$$

$$= \sin^{-1}\left(\frac{J_c}{\sqrt{J_c^2 - J_m^2}} \cos\theta \right) + \frac{\partial}{\partial J_c} \int \sqrt{2m\left[E - V(r)\right] - \frac{J_c^2}{r^2}} \, dr, \qquad (5.66)$$

$$\omega_r = \frac{\partial}{\partial J_r} \int \sqrt{2m\left[E - V(r)\right] - J_c^2/r^2} \, dr. \qquad (5.67)$$

Equation (5?67) may in principle be solved to express r in terms of J_c, J_r, and ω_r:

$$r = f_r(\omega_r, J_c, J_m) \qquad (5.68)$$

This expression for r may then be substituted into the second term of (5.66) yielding a relation of the form

$$\omega_c = \sin^{-1}\left(\frac{J_c}{\sqrt{J_c^2 - J_m^2}} \cos\theta \right) + f_c(\omega_r, J_c, J_m) \qquad (5.69)$$

The functions f_r and f_c will be periodic in ω_r, although they will in general contain all harmonics of the frequency ω_r.

Hence there is no restriction in the change of the quantum number n during a radiative transition. That is,

$$\Delta n = 0, \pm 1, \pm 2, \ldots \tag{5.70}$$

Exact expressions for the Fourier amplitudes of the functions f_r and f_e are obtained in Born, The Mechanics of the Atom, pp. 139-147, for the case of Kepler motion.

Introducing the angle α (see (1.227)),, where

$$\cos\alpha = \frac{J_m}{J_\ell}, \tag{5.71}$$

we may now write

$$\omega_m = \varphi - \sin^{-1}(\cot\alpha \cot\theta), \tag{5.72}$$

$$\omega_e = \sin^{-1}\left(\frac{\cos\theta}{\sin\alpha}\right) + f_e . \tag{5.73}$$

Hence

$$\cos\theta = \sin\alpha \sin(\omega_e - f_e), \tag{5.74}$$

$$\cot\theta = \frac{\sin\alpha \sin(\omega_e - f_e)}{\sqrt{1-\sin^2\alpha \sin^2(\omega_e - f_e)}}, \tag{5.75}$$

$$1 - \cot^2\alpha \cot^2\theta = \frac{1-\sin^2(\omega_e - f_e)}{1-\sin^2\alpha \sin^2(\omega_e - f_e)} = \frac{\cos^2(\omega_e - f_e)}{\sin^2\theta} \tag{5.76}$$

and finally

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ &= r\sin\theta\cos(\omega_m + \sin^{-1}(\cot\alpha\cot\theta)) \\ &= r\sin\theta\{\cos\omega_m\sqrt{1-\cot^2\alpha\cot^2\theta} - \sin\omega_m\cot\alpha\cot\theta\} \end{aligned}$$

$$= f_n \{\cos\omega_m \cos(\omega_e t - \phi_e) - \cos\alpha \sin\omega_m \sin(\omega_e t - \phi_e)\} \qquad (5.77)$$

$$\sqrt{1 - \cot^2\alpha\cot^2\theta} + \cos\omega_m \cot\alpha \cot\theta\}$$

$$= f_n \{\sin\omega_m \cos(\omega_e t - \phi_e) + \cos\alpha \cos\omega_m \sin(\omega_e t - \phi_e)\}, \qquad (5.78)$$

$$z = r\cos\theta$$

$$= f_n \sin\alpha \sin(\omega_e t - \phi_e). \qquad (5.79)$$

It is now seen that the components of the electric dipole moment (i.e. e_x, e_y, e_z) involve only the fundamental angular frequencies. ω_m, ω_e, without any of their higher harmonics. There are no constant terms in ω_e, although the z component of the electric dipole moment is constant in ω_m. Hence we have the selection rules

$$\Delta m = 0, \pm 1, \qquad (5.80)$$

$$\Delta l = \pm 1. \qquad (5.81)$$

A very large amount of work has been done in finding selection rules for many different quantum mechanical systems. Group theoretical methods have been extensively employed in this research. Since the study of selection rules is a very specialized field, we shall not consider it further here.

Thus far we have considered only the magnitudes of the energy jumps and the angular frequencies involved in radiative tran-

sitions. We have seen that the energy transfer process takes place by means of discrete units or quanta, but we have not examined the mechanics of the transfer process in any detail. Classically, this transfer process is continuous. In an attempt to retain as many classical ideas as possible we might try to suppose that, quantum mechanically, this process is also continuous -that is, that one system transfers energy to another in a continuous fashion just as in classical theory, the only difference now being that nothing observable happens until a full energy quantum has been exchanged. This hypothesis has been shown to be false by many experiments. For example, according to this hypothesis, electrons produced in a photoelectric process by very weak light should not begin to make their appearance until enough light has been absorbed to produce at least one full quantum. Actually, however, the electron production is observed to begin instantaneously, the only effect of weaker light being to produce fewer electrons. The intensity of the light, moreover, has nothing to do with the kinetic energy which the photoelectrons are observed to possess.

All this means that the energy transfer process is not classical in any respect save that of energy conservation. Application of the Correspondence Principle to this state of affairs forces us to the conclusion that quantum laws are fundamentally statistical in nature, as opposed to classical laws which are all deterministic. The quantum transfer process is an indivisible elementary process. All that the quantum theory can tell us is the <u>probability</u> of the process. Thus equation (5.44), instead of being understood as giving the rate of an energy radiation by a multiply periodic system, should be reinterpreted as giving the probable rate of emission of quanta from the system. Confining our attention to the transition from the m th state to the n th state, we may write

$$\left(\frac{dE^{rad}}{dt}\right)_{nm} = \frac{4}{3c^3}(\omega_{nm})^4 |\eta_{nm}|^2 \qquad (5.82)$$

This equation is to be understood as giving the <u>probable</u> rate of energy

radiation accompanied by a transition from the m th state to th n th. Since the angular frequency of the associated radiation is ω_{nm} energy radiation in amount $\hbar|\omega_{nm}|$ corresponds to the emission of one light quantum. Hence, we have for the probable rate of emission of light quanta, corresponding to the transition $m \longrightarrow n$,

$$R_{nm}^{emi} = \frac{1}{\hbar|\omega_{nm}|}\left(\frac{dE^{rad}}{dt}\right)_{nm} = \frac{4}{3c^3}\frac{|\omega_{nm}|^3}{\hbar}|p_{nm}|^2 \quad (5.83)$$

The quantum theory thus tells us only the rate at which quanta will probably be emitted. It does not tell us when any particular quantum will be emitted. Moreover, since quantum transfers are elementary indivible processes, it cannot describe the transition of the system from one quantized "orbit" to another. In the classical limit of high energies, however, the quantum numbers are large, many quanta are involved in transfer processes, and a single individual quantum transfer has an almost negligible effect, comparatively speaking, on the coupled systems, so that, because of the statistical nature of the individual transfers, the over-all process takes on a continuous and classical appearance. Even the "orbit" of the system becomes classically describable. The Correspondence Principle is therefore again satisfied.

The experimental fact that the quantum transfer process is an indivisible elementary process requires that all systems, radiation as well as matter, be subject to quantum laws and to their consequences such as energy quantization and statistical behavior. We have not up to this point undertaken the quantization of the electromagnetic field, nor shall we do so here. Quantization of the electromagnetic field would take us into the study of the general quantum theory of fields, a project upon which we not propose to embark at this time. It is a remarkable fact that many explicit formulas for processes involving emission and absorption of radiation may be inferred without actually going into quantum field theory. All we need to remember is that the radiation quanta are indivisible, that they have the simple

relation (5.1) between energy and frequency, and that they satisfy the Correspondence Principle in the classical limit. Radiation quanta can have the appearance of particles, when the measuring apparatus, by means of which they are observed, is designed so as to make them behave like particles (e.g. in the photoelectric and Compton effects). When they behave like particles they are commonly referred to as photons: Photons obey statistical laws. One cannot predict that a photon will definitely be at a certain place at a certain time, but in a beam of electromagnetic radiation there is a certain probability that a photon will cross a unit area in a unit of time. In the classical limit of many quanta, t is probability becomes proportional to the classical intensity of the beam.

In the discussion of the quantum radiation formulas (5.44), (5.82), (5.83) we have not considered the effect of radiation damping. We have seen in chapter 3 that, calssically, the process of radiation produces a damping force on the radiating system, which tends to diminish the amplitudes of the multiply periodic motion. Although in the quantum theory we cannot speak of a gradually diminishing amplitude, the classical damping effect has its quantum analogue. We shall first discuss the classical effet, beginning with the harmonic oscillator.

Referring to Eq. (3.155), we see that the equation of motion of a non-relativistic radiating harmonic oscillator is

$$m\ddot{x} - \frac{2}{3}\frac{e^2}{c^3}\dddot{x} + m\omega_0^2 x = 0 \tag{5.84}$$

or

$$\ddot{x} - \gamma^{-1}\dddot{x} + \omega_0^2 x = 0, \tag{5.85}$$

where the constant γ is given by (3.138) and where ω_0 is the oscillator angular frequency. Trying a complex solution of the form

$$x = x_0 e^{\lambda t} \tag{5.86}$$

we obtain the equation

$$\lambda^2 - \gamma^{-1}\lambda^3 + \omega_0^2 = 0. \qquad (5.87)$$

Equation (5.87) actually has three roots, corresponding to the approximate factorization

$$(\gamma^{-1}\lambda - 1)\left(\lambda - i\omega_0 + \tfrac{1}{2}\gamma^{-1}\omega_0^2\right)\left(\lambda + i\omega_0 + \tfrac{1}{2}\gamma^{-1}\omega_0^2\right)$$
$$\approx (\gamma^{-1}\lambda - 1)(\lambda^2 + \omega_0^2 + \gamma^{-1}\omega_0^2\lambda)$$
$$= \gamma^{-1}\lambda^3 - \lambda^2 - \omega_0^2.$$

Two of the roots are given by (5.88). The third is

$$\lambda = \gamma$$

leading to a "runaway" or "self-accelerated" solution

$$x = x_0 e^{\gamma t}$$

which must be disallowed.

Since γ^{-1} is normally a very small constant such that $\gamma^{-1}\omega_0 \ll 1$ (e.g. if e and m are the charge and mass of an electron), the following is, to excellent approximation, a solution of (5.87) :

$$\lambda = \pm i\omega_0 - \frac{1}{2}\gamma^{-1}\omega_0^2 \qquad (5.88)$$

We therefore have, as the real solution of (5.84),

$$x = x_0 e^{-\frac{1}{2}\gamma^{-1}\omega_0^2 t}\cos\omega_0 t. \qquad (5.89)$$

The amplitude of oscillation is seen to diminish exponentially.

The electric field at any point will vary in essentially the same way as x (neglecting small terms of order γ^{-1}).

$$E = E_0 e^{-\frac{1}{2}\gamma^{-1}\omega_0^2 t} \cos \omega_0 t. \tag{5.90}$$

Here t represents the retarded time at the point in question. It will now be convenient to make a Fourier analysis of this field. If the oscillator is assumed to be suddenly excited so that it begins radiating only at t = 0, and so that E = 0 prior to t = 0, then we may write

$$E(t) = \int_{-\infty}^{\infty} E(\omega)\, e^{i\omega t}\, d\omega \tag{5.91}$$

where

$$E(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} E(t)\, e^{-i\omega t}\, dt$$

$$= \frac{E_0}{4\pi}\int_{-\infty}^{\infty}\left[e^{i(\omega_0-\omega)-\frac{1}{2}\gamma^{-1}\omega_0^2]t} + e^{[-i(\omega_0+\omega)-\frac{1}{2}\gamma^{-1}\omega_0^2]t}\right] dt$$

$$= -\frac{E_0}{4\pi}\left[\frac{1}{i(\omega_0-\omega)-\frac{1}{2}\gamma^{-1}\omega_0^2} + \frac{1}{-i(\omega_0+\omega)-\frac{1}{2}\gamma^{-1}\omega_0^2}\right]. \tag{5.92}$$

The radiation intensity at the point in question, as given by the energy flux vector (see (3.176)) is

$$S = \frac{c}{4\pi} E^2 = \frac{c}{4\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(\omega)E(\omega')e^{i(\omega+\omega')t}\, d\omega\, d\omega' \tag{5.93}$$

Its time average value may evidently be expressed as

$$\bar{S} = \frac{1}{T}\int_0^{\infty} S(\omega)\, d\omega \tag{5.94}$$

where T is the "length" of the time axis and

$$S(\omega) = c\,|E(\omega)|^2. \tag{5.95}$$

$S(\omega)\, d\omega$ is the flux of radiation energy in the angular frequency

range ω to $\omega + d\omega$. We shall be interested in the distribution of this energy flux in the range near ω_0. In this range the second term in the brackets in (5.92) becomes negligible compared to the first, and we have

$$S(\omega) = \frac{c E_0^2}{16\pi^2} \frac{1}{(\omega_0 - \omega)^2 + \frac{1}{4}\gamma^{-2}\omega_0^4} \qquad (5.96)$$

The function $S(\omega)$ actually gives the intensity distribution in the spectral line emitted by the oscillator. When $|\omega_0 - \omega| = \frac{1}{2}\gamma^{-1}\omega_0^2$ the intensity is half of its maximum value. Hence the quantity $\gamma^{-1}\omega_0^2$ is called the "natural line breadth" of the harmonic oscillator.

Spectral lines in the quantum theory also possess a natural line breadth, which may be calculated by making appeal to the Correspondence Principle. Suppose that at time $t = 0$, a multiply periodic system is known to be in the m th energy state. Let Γ_m be the initial total rate of radiative transition out of this state to lower energy states. We have

$$\Gamma_m = \sum_{n<m} R_{n,m}^{emi} \qquad (5.97)$$

where $R_{n,m}^{emi}$ is given by (5.83). Let $P_m(t)$ denote the probability that the system is still in the m th state at a later time t. The rate of transition out of the m th state at any time must be proportional to the probability that the system is still in this state. Hence we may write

$$\dot{P}_m = -\Gamma_m P_m \quad \text{or} \quad P_m(t) = e^{-\Gamma_m t} \qquad (5.98)$$

Next, let $\omega_m(t)\,dt$ denote the probability that the system will make a transition from the m th state to some other state during the time interval t to t + dt. Then

$$P_m(t) = 1 - \int_0^t \omega_m(t)\,dt \qquad (5.99)$$

Differentiating (5.99), we obtain

$$\omega_m(t) = -P_m(t) = \Gamma_m e^{-\Gamma_m t} \qquad (5.100)$$

The average length of time which the particle takes to make the transition is therefore

$$T_m = \int_0^\infty t\,\omega_m(t)\,dt = \Gamma_m^{-1} \qquad (5.101)$$

Now, consider a specific transition, from the m th state to the n th state. The system will take, on the average, a length of time T_m to make this transition regardless of which lower energy state the n th may be. Although we cannot describe classically the quantum transition process, in the limit of large initial energies we must be able to give a classical description of the emitted radiation. We know that this radiation is "switched on" only during the length of time t', roughly equal to T_m, which it takes the system to make the transition. We don't know exactly which of the possible angular frequencies $\frac{E_m - E_n}{\hbar}$ the emitted quantum will choose to take, but we do know that once the quantum has settled its mind on a particular level, n , the emitted electric field must have the time behavior:

$$E = E_0 \cos|\omega_{mn}|t \qquad (5.102)$$

over a period of time t'. In (5.102) we have <u>neglected</u> the radiation damping, since at the higher energies it will have a negligible effect on the amplitude E_0 over a length of time T_m. The Fourier amplitudes of the electric field must therefore be given by

$$E_t(\omega) = \frac{1}{2\pi}\int_0^{t'} E(t)e^{-i\omega t}\,dt$$

$$= -\frac{E_0}{4\pi}\left[\frac{e^{-i(\omega-|\omega_{mn}|)t'}-1}{i(\omega-|\omega_{mn}|)} + \frac{e^{-i(\omega+|\omega_{mn}|)t'}-1}{i(\omega+|\omega_{mn}|)}\right] \qquad (5.103)$$

Confining our attention to the angular frequency range near $|\omega_{mn}|$, we have, using (5.95),

$$S_{t'}(\omega) = \frac{cE_0^2}{8\pi^2} \frac{1-\cos(|\omega_{mn}|-\omega)t'}{(|\omega_{mn}|-\omega)^2} \qquad (5.104)$$

In order to obtain the actually observed intensity distribution we must average this expression over the probability function $\omega_m(t')$.

$$S(\omega) = \int_0^\infty \omega_m(t') S_{t'}(\omega)\, dt'$$

$$= \frac{cE_0^2}{16\pi^3} \frac{\Gamma_m}{(|\omega_{mn}|-\omega)^2} \int_0^\infty e^{-\Gamma_m t'} [1-\cos(|\omega_{mn}|-\omega)t']\, dt'$$

$$= \frac{cE_0^2}{16\pi^3} \frac{\Gamma_m}{(|\omega_{mn}|-\omega)^2} \left[\frac{1}{\Gamma_m} - \frac{\Gamma_m}{(|\omega_{mn}|-\omega)^2+\Gamma_m^2}\right]$$

$$= \frac{cE_0^2}{8\pi^2} \frac{1}{(|\omega_{mn}|-\omega)^2+\Gamma_m^2} \qquad (5.105)$$

The natural breadth of the emitted line is thus seen to be equal to $2\Gamma_m$. For large energies $\Gamma_m \approx \Gamma_n$, so that the line breadth is also approximately equal to $2\Gamma_n$. But as we go to lower energies we should expect that the quantum line breadth Γ_{nm} for the transition $m \longrightarrow n$ would take some intermediate value between $2\Gamma_m$ and $2\Gamma_n$. Thus

$$\Gamma_{nm} \approx \frac{1}{2}(2\Gamma_m + 2\Gamma_n) = \Gamma_m + \Gamma_n \qquad (5.106)$$

Equation (5.106) is actually exact, as may be shown by the more accurate quantum mechanics to be developed in subsequent chapters. As an example, let us apply this equation to the transition $1 \longrightarrow 0$ for the harmonic oscillator. We then have $\Gamma_0 = 0$, and using Eqs. (5.48), (5.83) and (5.97)

$$\Gamma_1 = R_{01}^{emi} = \underbrace{\frac{4}{3c^3}\frac{\omega_0^3}{\hbar}\frac{e^2\hbar}{2m\omega_0}}_{=\gamma} |\eta_{01}|^2 = \gamma^{-1}\omega_0^2 \qquad (5.107)$$

That is, the line breadth is a consequence purely of radiation damping. The quantum line breadth, on the other hand, although obviously connected with the damping phenomenon, is more complicated. Only in the case of the transition $1 \rightarrow 0$ for the harmonic oscillator are the quantum and classical line breadths the same. This is connected with the fact that only for this transition does the oscillator come all the way to rest. The line breadths are broader for the higher transitions. At first sight this seems to contradict the Correspondence Principle. For as we go to higher energies we should expect a gradual approach to the classical situation. However, there is really no contradiction. To obtain the classical situation we must allow the quantized oscillator to make the whole series of transitions from a high energy level E_m to the ground level E_0. During this process, moreover, we must not interfere with the oscillator by making any measurements which might tell us which level the oscillator has reached at a given time. We must make simply an overall measurement of the intensity distribution of the emitted light. Quanta arising from the successive transitions will then be emitted <u>coherently</u> and will interfere with each other in such a way as to cancel out the broad tails of the individual intensity curves and reduce the whole distribution to the classical one.

The natural line breadth can evidently be described as an "uncertainty" ΔE_{nm} in the energy of the emitted photon :

$$\Delta E_{nm} = \hbar \Gamma_{nm} \qquad (5.108)$$

This, in turn, may be described to uncertainties in the energies of the two levels m and n :

$$\Delta E_{nm} = \Delta E_m + \Delta E_n \qquad (5.109)$$

where $$\Delta E_m = \hbar \Gamma_m, \quad \Delta E_n = \hbar \Gamma_n \tag{5.110}$$

The uncertainty ΔE_m in the energy of a given level is directly connected with the amount of time Δt_m which is available in which to measure this energy. This time is simply the probable lifetime of the level.

$$\Delta t_m = \tau_m \tag{5.111}$$

Combining now Eqs. (5.101), (5.110) and (5.111), we obtain the "uncertainty relation"

$$\Delta E_m \, \Delta t_m = \hbar \tag{5.112}$$

To summarize the preceding discussion, we find that the energy levels of a multiply periodic system are not really infinitely sharp as we had originally conceived them to be. Only the lowest level is infinitely sharp, because no radiation can take place from that level. If there were no coupling with the electromagnetic field, all levels would be infinitely sharp, but then, of course, we could not measure them. The lack of complete sharpness in the energy levels reflects itself in the lack of complete sharpness in the spectral lines of the emitted radiation. We find, rather, a continuous intensity distribution which, as a function of emitted frequency is sharply peaked around the quantum frequency calculated on the basis of no coupling with the electromagnetic field. The total integrated intensity in the neighborhood of the theoretical sharp line is proportional to the rate of emission of quanta corresponding to this line, that is R_{nni}^{emi}

The individual line breadths Γ_{nm} are actually quite small, being of the order of magnitude of a few times $\sigma^{-1}|\omega_{nm}|^2$. Hence, in practice, spectral lines still appear to be quite sharp. In the typical case in which the multiply periodic system is a mole-

cule or atom in a gas, there are a number of other phenomena which can produce a broadening of the spectral lines which completely overshadows the natural line breadth. Prominent among these there is the phenomenon known as "collision broadening". The line breadth due to collisions between the molecules or atoms can be calculated in exactly the same way as the natural line breadth. If Γ^{-1} is the average time between collisions then the probability that a particular molecule (or atom) has not yet experienced a collision after a time interval t is $e^{-\Gamma t}$ and the probability that it will suffer a collision between the times t and $t + dt$ is $\Gamma e^{-\Gamma t} dt$. Photons can be emitted coherently only during the times between collisions. Hence the Fourier analysis of the emitted radiation must have the form (5.103) and the line intensity distribution must have the form (5.105) with Γ_m replaced by Γ. The line breadth due to collision broadening is therefore 2Γ.

Having fairly discussed the basic principles of the emission of electromagnetic radiation from multiply periodic systems, we turn our attention now to the absorption of radiation by these systems. The action of an externally impressed radiation field upon a charged multiply periodic system may be described by a perturbation term H_1 in the total Hamiltonian function, given, in the non-relativistic approximation, by (see (4.19))

$$H_1 = -\eta \cdot E \tag{5.113}$$

where E is the electric vector and η the electric dipole moment. It is assumed that the dimensions of the system are small enough so that quadrupole and higher multiple forces are negligible. Referring to (5.42), we see that the amplitudes of the Fourier expansion (2.36) of H_1, are given by

$$H_{1\tau} = -\eta_\tau \cdot E. \tag{5.114}$$

We now imagine that the system is a member of an ensemble of identical systems oscillating with random phases. According to Eqs. (2.59, 60), the average energy absorbed per member of the ensemble over a long period of time T (i.e. long compared to the periods of oscillation) is then

$$\Delta E^{abs} = 4\pi^2 \sum_{\tau\omega>0} \tau_i \frac{\partial}{\partial \gamma_i} \left(\tau\omega \overline{|\eta_\tau \cdot E(\tau\omega)|^2} \right) \quad (5.115)$$

where

$$E(\tau\omega) = \frac{1}{2\pi} \int_{-T/2}^{T/2} E(t)\, e^{-i\tau\omega t}\, dt \quad (5.116)$$

and where the bar denotes an average taken over all the systems, it being remembered that both η_τ and $E(t)$, and hence $\eta_{\tau c}$ and $E(\omega)$ may vary from one system to another in space. If the radiation is assumed to be isotropic and the systems are distributed and oriented at random, then

$$\overline{\eta_{\tau i} \eta_{-\tau j}} = \frac{1}{3} \delta_{ij} |\eta_\tau|^2, \quad (5.117)$$

$$\overline{E_i(\tau\omega) E_j(-\tau\omega)} = \frac{1}{3} \delta_{ij} |E(\tau\omega)|^2, \quad (5.118)$$

and

$$\overline{|\eta_\tau \cdot E(\tau\omega)|^2} = \overline{\eta_{\tau i} \eta_{-\tau j} E_i(\tau\omega) E_j(-\tau\omega)}$$

$$= \overline{\eta_{\tau i} \eta_{-\tau j}}\ \overline{E_i(\tau\omega) E_j(-\tau\omega)}$$

$$= \frac{1}{3} |\eta_\tau|^2 |E(\tau\omega)|^2, \quad (5.119)$$

Hence

$$\Delta E^{abs} = \frac{4\pi^2}{3} \sum_{\tau\omega>0} \tau_i \frac{\partial}{\partial y_i} \left(\tau\omega |\eta_\tau|^2 |E(\tau\omega)|^2 \right). \quad (5.120)$$

Equation (5.120) may be reexpressed in terms of the energy distribution function of the radiation field. The energy density in the field (see (3.174)) is

$$U = \frac{E^2}{4\pi} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{E}(\omega) \cdot E(\omega') e^{i(\omega+\omega')t} d\omega d\omega', \quad (5.121)$$

Its time average value is therefore

$$\bar{U} = \int_0^\infty U(\omega) d\omega \quad (5.122)$$

where

$$U(\omega) = \frac{1}{T} |E(\omega)|^2 \quad (5.123)$$

and where

$$E(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt. \quad (5.124)$$

Eq. (5.120) thus becomes

$$\frac{\Delta E^{abs}}{T} = \frac{4\pi^2}{3} \sum_{\tau\omega>0} \tau_i \frac{\partial}{\partial y_i} \left(\tau\omega |\eta_\tau|^2 U(\tau\omega) \right). \quad (5.125)$$

Equation (5.125) gives the classical average rate of absorption of radiant energy by members of the ensemble. By making use of the Correspondence Principle we can also write down its quantum analogue. We must first find the quantum analogue of the partial derivative $\partial/\partial y_i$. Suppose the typical system is in the m th energy state. We have already become familiar with the correspondence

$$\tau_i \omega^i = \tau_i \frac{\partial E}{\partial J_i} \longleftrightarrow \omega_{nm} = \frac{E_n - E_m}{\hbar} \qquad (5.126)$$

with

$$n_i - m_i = \tau_i \qquad (5.127)$$

in which a derivative in the classical theory is replaced by a difference in the quantum theory. Following the same idea, therefore, we write

$$\sum_{\tau\omega > 0} \tau_i \frac{\partial}{\partial J_i} \left(\tau\omega |\eta_\tau|^2 U(\tau\omega) \right)$$

$$\longleftrightarrow \sum_{\substack{n' \\ \tau_i = n_i' - m_i' \\ E_{n'} > E_{m'}}} \frac{1}{\hbar} \left\{ \left(\tau\omega |\eta_\tau|^2 U(\tau\omega) \right)(J_{n'}) - \left(\tau\omega |\eta_\tau|^2 U(\tau\omega) \right)(J_{m'}) \right\}$$

$$\longleftrightarrow \sum_{\substack{n \\ E_n > E_m \\ n_i - m_i = n_i' - m_i' = n_i - n_i''}} \frac{1}{\hbar} \left\{ \omega_{nm} |\eta_{nm}|^2 U(\omega_{nm}) - \omega_{nn''} |\eta_{mn''}|^2 U(\omega_{nn''}) \right\} \qquad (5.128)$$

where n_i' and m_i' are quantum numbers which roughly straddle the quantum number n_i. In the final sum in (5.128) the number n_i takes on larger and large values while the number n_i'' takes on smaller and smaller values in such a way that $n_i - m_i = n_i - n_i''$ for all i. Strictly speaking, of course, this is not quite possible since, although n_i can increase without limit, n_i'' cannot go below a certain minimum value, usually zero. However, in the classical limit of high quantum numbers the neglect of terms for which n_i'' would be negative is unimportant. Therefore, remembering that $\omega_{nm} = -\omega_{mn}$, $\eta_{nm} = \eta_{mn}^*$, we may combine the two terms inside the brackets in (5.128) into a single summation and write

$$\frac{\Delta E^{abs}}{T} = \frac{4\pi^2}{3\hbar} \sum_n \omega_{nm} |\eta_{nm}|^2 U(|\omega_{nm}|) \qquad (5.129)$$

Equation (5.129) is a special case of the quantum analogue of the more general classical equation (2.59), namely

$$\Delta E^{abs} = \frac{1}{\hbar} \sum_n \omega_{nm} |h_{1nm}|^2. \qquad (5.130)$$

The rate at which the oscillating systems absorb energy from the radiation field in passing from the m th energy state to the n th is evidently given by

$$\left(\frac{dE^{abs}}{dt}\right)_{nm} = \frac{4\pi^2}{3\hbar}\omega_{nm}|\eta_{nm}|^2 U(|\omega_{nm}|) \quad (5.131)$$

This rate is seen to be proportional to the energy density of the radiation field at the angular frequency $|\omega_{nm}|$. Moreover, when $E_n < E_m$ this rate is negative. That is, for $E_n < E_m$ the sytem <u>omits</u> energy rather than absorbs it. This process is known as "induced emission". The rate of induced emission in passing from an energy state m to a lower energy state n is seen to be equal to the rate of induced absorption in passing from the state n to the state m .

We may readily use the preceding results to derive Planck's well known radiation formula. Since, however, this is really a formula of statistical mechanics, we shall first briefly review the steps in deriving the well known equilibrium distribution formulae. Consider a macroscopic system composed of a very large number N of subsystems, i.e. an ensemble. The subsystems may be supposed to be multiply periodic. Consider a typical subsystem. Suppose that its quantum states are in some convenient way divided up into groups labeled by indices i, j etc , the states in each group having roughly equal energy. Let E_i denote the energy of the i th group and let ρ_i denote the number of states in the i th group. Suppose that a "macroscopic state" of the ensemble is determined by specifying for all i the number n_i of subsystems to be found having quantum states in the i th group. We seek the number $N(n_1, n_2 \ldots)$ of distinct quantum states of the ensemble which give rise to one and the same macroscopic state $n_1, n_2, n_3 \ldots$.

Suppose first that the subsystems are all distinguishable. Then the quantum state of the ensemble is specified by giving the quantum states of all of the subsystems. The number $N(n_1, n_2 \ldots)$ is there-

fore equal to the number of ways in which N distinguishable objects can be arranged into a certain number of groups i, multiplied by the number of ways the objects can be placed in cells within each group, there being g_i cells in the i th group. We have

$$N(n_1, n_2 \ldots) = N! \prod_i \frac{g_i^{n_i}}{n_i!} \tag{5.132}$$

$N_{n_1 n_2 \ldots}$ is in general a very peaked function of the variables n_i. Hence for any function $f(n_1, n_2, \ldots)$ of the n_i

$$\int f(n_1, n_2 \ldots) N(n_1, n_2 \ldots) dn_1 dn_2 = f(n_1^{max}, n_2^{max} \ldots) \tag{5.133}$$

where n_i^{max} are the values of the variables n_i which make $N(n_1 n_2 \ldots)$ a maximum. If all quantum states are regarded as occurring with equal probability in time (there are always weak interactions which cause transitions from state to state) then the integral in (5.133) represents the mean or probable value of the function f. We write

$$\overline{f} = f(n_1^{max}, n_2^{max} \ldots) \tag{5.134}$$

In particular

$$\overline{n_i} = n_i^{max} \tag{5.135}$$

In order to maximize $N(n_1, n_2 \ldots)$ we differentiate Eq. (5.132) or rather its logarithm

$$\log N(n_1, n_2 \ldots) = -\sum_i (n_i \log n_i - n_i \log g_i) + Const. \tag{5.136}$$

Here we have assumed the n_i to be very large numbers and have neglected terms of smaller order of magnitude. Remembering the restrictions

$$\sum_i n_i = N, \quad \sum_i n_i E_i = E \tag{5.137}$$

where E is the total energy of the ensemble, we have, for a variation δn_i at maximum,

$$0 = -\delta \sum_i (n_i \log n_i - n_i \log g_i + \alpha n_i + \beta n_i E_i)$$

$$= -\sum_i (\log n_i - \log g_i + \alpha + \beta E_i)\delta n_i, \qquad (5.138)$$

where α and β are multipliers. Hence

$$\bar{n}_i = n_i^{max} = g_i A e^{-E_i/kT} \qquad (5.139)$$

where $$A = e^{-\alpha}, \quad kT = \frac{1}{\beta}. \qquad (5.140)$$

k and T may be experimentally identified as Boltzmann's constant and the absolute temperature of the ensemble respectively. Equation (5.139) is known as the Boltzmann distribution formula, and is applicable when the subsystems of the ensemble are distinguishable. Distinguishability of subsystems is known as Boltzmann statistics.

Applying Eq. (5.139) to our ensemble of oscillators interacting with the radiation field, and letting the groups i coincide with the individual subsystem states, we have $g_i = 1$, and

$$N_m = A e^{-E_m/kT}, \quad N_n = A e^{-E_n/kT} \qquad (5.141)$$

where N_m and N_n are respectively the probable numbers of oscillators in the m th and n th states. Suppose $E_n > E_m$. Then the transition $m \rightarrow n$ is an absorbtive transition governed by Eq. (5.131). The total average rate at which oscillators in the ensemble undergo the transition $m \rightarrow n$ is

$$R_{nm} = N_m \frac{1}{\hbar \omega_{nm}} \left(\frac{dE^{abs}}{dt}\right)_{nm}$$

$$= A \frac{4\pi^2}{3h^2} |\eta_{nm}|^2 e^{-E_m/kT} U(|\omega_{nm}|). \qquad (5.142)$$

The transition $n \rightarrow m$, on the other hand, is an emissive transition, governed partly by Eq. (5.83) which describes spontaneous emission and partly by Eq. (5.131) describing induced emission. The total average rate at which oscillators in the ensemble undergo the transition $n \longrightarrow m$ is

$$R_{mn} = N_n \left[R_{mn}^{emi} - \frac{1}{\hbar\omega_{nm}} \left(\frac{dE^{abs}}{dt} \right)_{mn} \right]$$

$$= A \left[\frac{4}{3c^3} \frac{|\omega_{nm}|^3}{\hbar} + \frac{4\pi^2}{3h^2} U(|\omega_{nm}|) \right] |\eta_{nm}|^2 e^{-\frac{E_m}{kT}}. \qquad (5.143)$$

Now suppose that all the possible angular frequencies are at least slightly different from one another. (If they are not initially different, the introduction of slight perturbations can make them so.) Then the emission or absorption of photons with angular frequency $|\omega_{nm}|$ can take place only by means of the transitions $m \leftrightarrow n$. When a condition of equilibrium is reached between the oscillators and the radiation, the detailed balancing condition must then be satisfied, namely

$$R_{mn} = R_{nm}. \qquad (5.144)$$

Combining equations (5.142, 143, 144), we obtain

$$\frac{4\pi^2}{3h^2} e^{\frac{E_n - E_m}{kT}} U(|\omega_{nm}|) = \frac{4}{3c^3} \frac{|\omega_{nm}|^3}{\hbar} + \frac{4\pi^2}{3h^2} U(|\omega_{nm}|) \qquad (5.145)$$

Equation (5.145) holds for all the angular frequencies $|\omega_{nm}|$. If we now suppose that the ensemble includes many different types of subsystems so that virtually all possible frequencies are represented, then we obtain the following general energy distribution function for the radiation (remembering $E_n - E_m = \hbar\omega_{nm}$):

$$U(\omega) = \frac{\frac{4}{c^3}\frac{\omega^3}{\hbar}}{\frac{4\pi^2}{3\hbar^2}\left(e^{\hbar\omega/kT} - 1\right)} = \frac{\hbar}{\pi^2 c^3}\,\frac{\omega^3}{e^{\hbar\omega/kT} - 1} \qquad (5.146)$$

Equation (5.146) is Planck's radiation formula. It is seen to be independent of the nature of the oscillators producing the radiation.

The lack of dependence of (5.146) on the nature of the oscillators may be demonstrated in a more direct fashion by means of a derivation which makes use of the properties of the radiation field alone. Consider a radiation field confined to a cubical box of volume $V = L^3$. This field can be regarded as a superposition of plane standing waves with propergation vectors $\vec{k}$ given by

$$k_i = \frac{n_i \pi}{L} \quad , \quad i = 1,2,3 \qquad (5.147)$$

where the n_i are positive intergers. The angular frequency of a particular standing wave is

$$\omega = c|k| = \frac{\pi c}{L} n \qquad (5.148)$$

where

$$n = \left(n_1^2 + n_2^2 + n_3^2\right)^{1/2}. \qquad (5.149)$$

Each standing wave may be regarded as a harmonic oscillator. The number of oscillators in the angular frequency range ω to $\omega + d\omega$ is evidently

$$2 \times \frac{1}{8} \times 4\pi^2 n\, dn = \frac{V}{\pi^2 c^3}\,\omega^2 d\omega \qquad (5.150)$$

The factor 2 corresponds to the fact that there are two polarization directions, and hence two oscillators, for each set of intergers n_i. The factor 1/8 refers to the fact that only one octant of n -space is

involved. The same result can be obtained by using a periodicity condition $k_i = \frac{2n_i\pi}{L}$ instead of the containment condition (5.147) and allowing the n_i to range over negative as well as positive values.

Using now the Boltzmann formula (5.139), we have, for the probable number of oscillators in the range ω to $\omega + d\omega$ which are to be found in the m th energy level,

$$dN_m = dA\, e^{-E_m/kT} = dA\, e^{-m\hbar\omega/kT}, \qquad (5.151)$$

where we have dropped the unobservable "zero point energy" of the quantized oscillators. dA is determined by the condition

$$\frac{V}{\pi^2 c^3}\omega^2 d\omega = \sum_{m=0}^{\infty} dN_m = dA \sum_{m=0}^{\infty} e^{-m\hbar\omega/kT}$$

$$= \frac{dA}{1 - e^{-\hbar\omega/kT}} \qquad (5.152)$$

The probable radiation energy density in the range ω to $\omega + d\omega$ is therefore

$$U(\omega)\,d\omega = \frac{1}{V}\sum_{m=0}^{\infty} E_m\, dN_m = \frac{\omega^2 d\omega}{\pi^2 c^3}\left(1 - e^{-\frac{\hbar\omega}{kT}}\right)\sum_{m=0}^{\infty} m\hbar\omega\, e^{-\frac{m\hbar\omega}{kT}}$$

$$= \frac{\hbar\omega^3 d\omega}{\pi^2 c^3}\,\frac{e^{-\frac{\hbar\omega}{kT}}}{1 - e^{-\frac{\hbar\omega}{kT}}} \qquad (5.153)$$

which leads to (5.146).

Instead of regarding each standing wave as a harmonic oscillator, we may regard it as being composed of a certain number of photons, the probable number of these photons being simply the mean value of the quantum number of the corresponding oscillator, namely

$$n_\omega = \frac{\sum_{m=0}^{\infty} m e^{-m\hbar\omega/kT}}{\sum_{m=0}^{\infty} e^{-m\hbar\omega/kT}} = \frac{1}{e^{\hbar\omega/kT} - 1}, \qquad (5.154)$$

In the interaction of the radiation field with a single material oscillator the probability or rate for the transition $m \longrightarrow n$ with the simultaneous absorption of one photon from a particular standing wave of angular frequency ω_{nm}, must be proportional to the mean number of photons which are present in this standing wave before the absorption process. The average rate at which the material oscillator makes the transition $m \longrightarrow n$ by absorbing a photon from any standing wave of angular frequency must be proportional to the density of standing waves at this frequency (see (5.150)). Denoting this rate by R_{nm}, we have therefore

$$R_{nm} = C_{nm} \frac{|\omega_{nm}|^2}{\pi^2 c^3} n_{\omega_{nm}} \qquad (5.155)$$

where C_{nm} is a certain proportionality constant.

The probability for the reverse transition, $n \longrightarrow m$, with the emission of a photon into the standing wave in question, may be inferred by appealing to the principle of reversibility of time. Remembering that the dynamical equations describing the interaction of matter and radiation are invariant under time reversal ($t \longrightarrow -t$, $e \longrightarrow -e$), that emission and absorption processes are independent of the sign of the charge (being proportional to $|\eta|^2$), and that an absorption process becomes, under reversal of time, an emission process, we see that this probability must be proportional to the mean number of photons which are present in this standing wave after the emission process has taken place. Thus, for the emission rate into all standing waves of angular frequency ω_{nm}, we have

$$R_{mn} = C_{nm} \frac{|\omega_{nm}|^2}{\pi^2 c^3} (n_{\omega_{nm}} + 1). \qquad (5.156)$$

In a macroscopic ensemble of interacting material and radiation oscillators we have, therefore, for the total average rate at

which material oscillators undergo the transitions $n \leftrightarrow m$,

$$R_{nm} = N_n r_{nm} = \frac{C_{nm}}{\hbar|\omega_{nm}|} A e^{-\frac{E_n}{kT}} U(|\omega_{nm}|), \quad (5.157)$$

$$R_{mn} = N_m r_{mn} = \frac{C_{nm}}{\hbar|\omega_{nm}|} A e^{-\frac{1}{kT}\left[U(|\omega_{nm}|) + \frac{\hbar|\omega_{nm}|^3}{\pi^2 c^3}\right]} \quad (5.158)$$

where we have used the notation of Eqs. (5.141-145) and the results of Eqs. (5.146), (5.153), (5.154). Equations (5.157, 158) are consistent with (5.142, 143) provided we set

$$C_{nm} = \frac{4\pi^2}{3\hbar} |\omega_{nm}| \, |\eta_{nm}|^2. \quad (5.159)$$

In the preceding discussion only the number of photons in each standing wave is important; the photons themselves being completely indistinguishable. This indistinguishability may be made use of in still another derivation of the Planck radiation formula. Consider the radiation field as an ensemble of photons instead of an ensemble of oscillators. The quantum state of each photon is determined by specifying the standing wave (i.e. the propagation vector k) with which the photon is associated. Since the photons are indistinguishable, the quantum state of the radiation field as a whole is completely specified by simply giving the number of photons associated with each permissable propagation vector k. In the preceding derivation of the Planck formula, Boltzmann statistics were applied to the individual distinguishable radiation oscillators. Owing to the indistinguishability of photons, however, a new kind of statistics will have to be introduced when we regard the photons rather than the oscillators as the fundamental units of the radiation field. These new statistics are known as Bose-Einstein statistics. The two characteristics of Bose-Einstein statistics are : 1) the indistinguishability of the ensemble subsystems, and 2) the fact that an unlimited number of subsystems can exist in the same subsystem quantum state.

We first derive the general equilibrium formula for Bose-Einstein statistics. As in the case of Boltzmann statistics we divide the subsystem quantum states into monoenergetic groups, and we assume that the macroscopic state is determined by specifying the number of subsystems in each group. Using the same notion as previously, we have, for the number of distinct quantum states which give rise to one and the same macroscopic state,

$$N(n_1, n_2, \ldots) = \prod_i \frac{(n_i + g_i - 1)!}{n_i!\,(g_i - 1)!} \qquad (5.160)$$

This number is simply the product over i of the number of ways in which n_i indistinguishable objects can be placed into g_i cells, there being no restriction on the number of objects per cell. This is also the product over i of the number of different arrangements of n_i dots and $g_i - 1$ lines in patterns of the form

$$\cdot|\ldots|\cdot\cdot\cdot|\cdot|\cdot\cdot\cdot\cdot\cdot|\cdot\cdot\cdot\cdot|\cdot|\cdot|\cdot|\cdot|\cdot|\cdot|\cdot\cdot|\cdot\cdot\cdot$$

Since photons can be emitted and absorbed their number is not fixed, so that we do not have the first of the constraints (5.137) The second of these constraints, however, is still valid, since the energy in the radiation field is assumed to be fixed. (It is actually determined, of course, by the temperature of the ensemble of matetiel oscillators with which it is in equilibrium.) The situation can best be clarified by including the material oscillators as well as the photons in the maximizing procedure which follows. Hence, writing

$$\log N(n_1, n_2 \ldots) = \sum_i \left[(n_i + g_i)\log(n_i + g_i) - n_i \log n_i\right] \qquad (5.161)$$

where we have assumed the n_i and g_i to be large numbers, we have, for a variation at maximum,

$$0 = \delta \sum_i \left[(n_i + g_i)\log(n_i + g_i) - n_i \log n_i - \beta n_i E_i\right]$$

$$= \sum_i \left[\log(n_i + g_i) - \log n_i - \beta E_i\right] \delta n_i \qquad (5.162)$$

where β is the same multiplier as in (5.138). Using (5.140), we have

$$\bar{n}_i = n_i^{max} = \frac{\rho_i}{e^{\epsilon_i/kT} - 1} \tag{5.163}$$

We now apply eq. (5.163) in differential form. The number, $d\rho$, of photon states corresponding to angular frequencies lying in the range ω to $\omega + d\omega$ is simply the number of standing waves in this range. This number is given by (5.150). Thus

$$d\rho = \frac{V}{\pi^2 c^3} \omega^2 d\omega. \tag{5.164}$$

The probable number of photons to be found in this range is, by (5.163)

$$dn = \frac{d\rho}{e^{\hbar\omega/kT} - 1} \tag{5.165}$$

The average photon energy density in this range is therefore

$$U(\omega)\, d\omega = \frac{1}{V} \hbar\omega\, dn = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\hbar\omega/kT} - 1} d\omega, \tag{5.166}$$

and we are once again led to the Planck formula.

In all the preceding discussion we have considered the process[23] of emission and absorption separately. We shall now turn to a consideration of situations in which absorption and emission processes take place simultaneously and in an inseparable fashion. Specifically, we shall consider the scattering of electromagnetic radiation by a multiply periodic charged material system. In a scattering process the physical picture is as follows. The incoming electromagnetic field interacts with the electric dipole moment of the system, causing it to oscillate in a different fashion from that of the unperturbed system. This alteration in the oscillation produces an alteration in the radiation which is emitted by the system. The unperturbed system normally undergoes a process of spontaneous emission of radiation if

it is initially in an excited state. This process we have already examined. The perturbed system, however, undergoes an additional emission process owing to the forced oscillations produced by the incoming radiation. The additional emitted radiation is coherent with the incoming radiation. The over-all-effect is that of scattering of the incident radiation by the multiply periodic system.

We first treat the problem calssically, using the time-dependent perturbation theory of chapter 2. We expand the dipole moment in terms of the angle and action variables of the unperturbed system (see (5.42)).

$$\eta = \sum_{\tau} \eta_{\tau}(J_0)\, e^{i\tau\omega_0} \tag{5.167}$$

In terms of the canonical variables ω^i, J_i of the perturbed system this becomes (cf. (2.51)).

$$\eta(\omega_i, J_0) = \eta(\omega, J) - (\eta, W_1)(\omega, J) + \dots$$
$$= \eta^{'0} + \eta^{'1} + \dots \tag{5.168}$$

where

$$\eta^{'0} = \sum_{\tau} \eta_{\tau}(J)\, e^{i\tau\omega} \tag{5.169}$$

and

$$\eta^{'1} = -\sum_{\tau,\tau'} (\eta_{\tau} e^{i\tau\omega}, W_{1\tau'} e^{i\tau'\omega})$$

$$= -\sum_{\tau,\tau'} \left[-i\tau_i \frac{\partial \eta_{\tau}}{\partial J_i} W_{1\tau'} + i\tau_i \eta_{\tau} \frac{\partial W_{1\tau'}}{\partial J_i} \right] e^{i(\tau+\tau')\omega} \tag{5.170}$$

The quantities inside the brackets of the last expression are functions of the J's.

The perturbation function has the form (see (5.113, 114)

$$H_1 = -\eta \cdot E = \sum_\tau H_{1\tau}(\gamma_0, t) e^{i\tau\omega_0} \tag{5.171}$$

where

$$H_{1\tau}(\gamma_0, t) = -\eta_\tau(\gamma_0) \cdot E(t). \tag{5.172}$$

Here E is the incoming electric field vector evaluated at a suitable point within the system. We shall take it to be that corresponding to a monochromatic plane of angular frequency ω.

$$E(t) = \ell(\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}). \tag{5.173}$$

ϵ is a complex amplitude and ℓ is a unit vector which designates the direction of polarization of the plane wave. Making use of (2.45), we have

$$W_{1\tau} = -\int_{-\infty}^{t} (-\ell \cdot \eta_\tau)(\epsilon e^{i\omega t'} + \epsilon^* e^{-i\omega t'}) e^{i\tau\omega_0(t'-t)} dt'$$

$$= \ell \cdot \eta_\tau e^{-i\tau\omega_0 t} \int_{-\infty}^{t} [\epsilon e^{i(\omega+\tau\omega_0)t'} + \epsilon^* e^{-i(\omega-\tau\omega_0)t'}] dt'$$

$$= \ell \cdot \eta_\tau \left[\epsilon \frac{e^{i\omega t}}{i(\omega+\tau\omega_0)} - \epsilon^* \frac{e^{-i\omega t}}{i(\omega-\tau\omega_0)}\right] \tag{5.174}$$

in which the value of the integral at the lower limit is taken to be zero, as if there were an exponential damping factor in the integrand corresponding to an adiabatic switching on of the incoming field.

In order to calculate the field radiated by the system we must known the time dependence of η. Since $\dot{\gamma}_i = 0$ (see (2.50)) this is almost given to us by (5.169, 170). We need to know only the

time dependence of the ω^i. Integrating equation (2.49) and using (5.172, 173), we find

$$\omega^i = \int \left(\omega_0{}^i + \frac{\partial H_{10}}{\partial J_i} + \dots\right) dt$$

$$= \omega_0{}^i t + \varphi_i - \frac{\partial \eta_2}{\partial J_i} \cdot l \int \left(\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}\right) dt + \dots$$

$$= \omega_0^i t + \varphi^i - \frac{\partial \eta_2}{\partial J_i} \cdot l \frac{1}{i\omega} \left(\epsilon e^{i\omega t} - \epsilon^* e^{-i\omega t}\right) + \dots \tag{5.175}$$

where the φ^i are arbitrary constant phases. Hence

$$\eta'^0 = \sum_\tau \eta_\tau e^{i\tau(\omega_0 t + \varphi)} e^{-\tau_i \frac{\partial \eta_2}{\partial J_i} l \frac{1}{\omega} (\epsilon e^{i\omega t} - \epsilon^* e^{-i\omega t}) + \dots}$$

$$= \sum_\tau \eta_\tau \left[1 - \tau_i \frac{\partial \eta_2}{\partial J_i} \cdot l \frac{1}{\omega} \left(\epsilon e^{i\omega t} - \epsilon^* e^{-i\omega t}\right)\right] e^{i\tau(\omega_0 t + \varphi)} + \dots \tag{5.176}$$

and

$$\eta'^1 = -\sum_{\tau,\tau'} \left[-i\tau_i \frac{\partial \eta_{\tau-\tau'}}{\partial J_i} W_{\tau'} + i(\tau_i - \tau_i) \eta_{\tau-\tau'} \frac{\partial W_{\tau'}}{\partial J_i}\right] e^{i\tau(\omega_0 t + \varphi)} + \dots$$

$$= \sum_\tau \sum_{\tau'} \left\{ i\tau_i \frac{\partial}{\partial J_i} \left[\eta_{\tau-\tau'} \cdot l \cdot \eta_{\tau'} \left(\epsilon \frac{e^{i\omega t}}{i(\omega + \tau'\omega_0)} - \epsilon^* \frac{e^{-i\omega t}}{i(\omega - \tau'\omega_0)}\right)\right]\right.$$

$$\left. - i\tau_i \eta_{\tau-\tau'} \frac{\partial}{\partial J_i} \left[l \cdot \eta_{\tau'} \left(\epsilon \frac{e^{i\omega t}}{i(\omega + \tau'\omega_0)} - \epsilon^* \frac{e^{i\omega t}}{i(\omega - \tau'\omega_0)}\right)\right]\right\} e^{i\tau(\omega_0 t + \varphi)}$$

$$+ \dots \tag{5.177}$$

Evidently we may write, correct to second order

$$\eta = \eta^0 + \eta^1 \tag{5.178}$$

where

$$\eta^{\circ} = \sum_{\tau} \eta_{\tau} e^{i\tau(\omega_0 t + \varphi)} \tag{5.179}$$

and

$$\eta' = \sum_{\tau,\tau'} \left\{ \tau_i \frac{\partial}{\partial j_i} \left[\varepsilon \frac{\eta_{\tau-\tau'} \ell \eta_{\tau'}}{\omega + \tau'\omega_0} \right] - \tau_i \eta_{\tau-\tau'} \frac{\partial}{\partial j_i} \left[\varepsilon \frac{\ell \eta_{\tau'}}{\omega - \tau'\omega_0} \right] \right\} e^{-i(\omega - \tau\omega_0)t + i\tau\varphi}$$

$$- \sum_{\tau,\tau'} \left\{ \tau_i \frac{\partial}{\partial j_i} \left[\varepsilon^* \frac{\eta_{\tau'-\tau} \ell \eta_{\tau'}}{\omega - \tau'\omega_0} \right] - \tau_i \eta_{\tau'-\tau} \frac{\partial}{\partial j_i} \left[\varepsilon^* \frac{\ell \eta_{\tau'}}{\omega - \tau'\omega_0} \right] \right\} e^{-i(\omega - \tau\omega_0)t + i\tau\varphi} \tag{5.180}$$

The component, η°, of the dipole moment vector gives rise to the ordinary spontaneously emitted radiation. η', on the other hand, accounts for the scattered radiation, and is the component which is of interest to us at present;

Let n be an arbitrary unit vector. According to the results of Problem XXV, the energy flux of the scattered radiation in the direction n and at a distance r from the scattering system is

$$S^{scat} = \frac{1}{4\pi c^3} \frac{(n \times \ddot{\eta}')^2}{r^2} \tag{5.181}$$

the dots denoting differentiation with respect to the time. In order to calculate the scattering cross section of the system we shall need the time average values of this expression. It is evident that only the cross product of the terms in (5.180) can contribute to this average unless $2\omega = \tau\omega_0$ for some τ_i. If we include an average over the phases, however, then any product terms for which $2\omega = \tau\omega_0$ are excluded, and only the cross product is left. Thus

$$\overline{(\eta \times \ddot{\eta}^{1})^{2}} = 2\varepsilon\varepsilon^{*} \sum_{\tau,\tau',\tau''} (\omega - \tau\omega_{0})^{4}$$

$$\times \left\{ \tau_{i}' \frac{\partial}{\partial \gamma_{i}} \left[\frac{n \times \eta_{-\tau+\tau'} \cdot \ell \cdot \eta_{-\tau'}}{\omega - \tau'\omega_{0}} \right] - \tau_{i} n \times \eta_{-\tau-\tau'} \frac{\partial}{\partial \gamma_{i}} \left[\frac{\ell \cdot \eta_{\tau'}}{\omega - \tau'\omega_{0}} \right] \right\}$$

$$\times \left\{ \tau_{i}'' \frac{\partial}{\partial \gamma_{i}} \left[\frac{n \times \eta_{-\tau-\tau''} \cdot \ell \cdot \eta_{\tau''}}{\omega - \tau''\omega_{0}} \right] - \tau_{i} n \times \eta_{\tau-\tau''} \frac{\partial}{\partial \gamma_{i}} \left[\frac{\ell \cdot \eta_{\tau''}}{\omega - \tau''\omega_{0}} \right] \right\}$$

$$= 2\varepsilon\varepsilon^{*} \sum_{\tau} (\omega - \tau\omega_{0})^{4} \Big/ \sum_{\tau'} \left\{ \tau_{i}' \frac{\partial}{\partial \gamma_{i}} \left[\frac{n \times \eta_{\tau-\tau'} \cdot \ell \cdot \eta_{\tau'}}{\omega - \tau'\omega_{0}} \right] - \tau_{i} n \times \eta_{\tau-\tau'} \frac{\partial}{\partial \gamma_{i}} \left[\frac{\ell \cdot \eta_{\tau'}}{\omega - \tau'\omega_{0}} \right] \right\}^{2} \tag{5.182}$$

The mean value of the energy flux of the incident radiation is

$$\overline{S^{in}} = \frac{c}{4\pi} \overline{E^{2}} = \frac{c}{2\pi} \varepsilon\varepsilon^{*} \tag{5.183}$$

and we have, for the differential scattering cross section,

$$\frac{d\sigma}{d\Omega} = \frac{r^{2} \overline{S_{scat}}}{\overline{S^{in}}}$$

$$= \frac{1}{c^{4}} \sum_{\tau} (\omega - \tau\omega_{0})^{4} \Big/ \sum_{\tau'} \left\{ \tau_{i}' \frac{\partial}{\partial \gamma_{i}} \left[\frac{n \times \eta_{\tau-\tau'} \cdot \ell \cdot \eta_{\tau'}}{\omega - \tau'\omega_{0}} \right] - \tau_{i} n \times \eta_{\tau-\tau'} \frac{\partial}{\partial \gamma_{i}} \left[\frac{\ell \cdot \eta_{\tau'}}{\omega - \tau'\omega_{0}} \right] \right\}^{2} \tag{5.184}$$

The quantum analogue of equation (5.184) may readily be obtained by use of the Correspondence Principle. One proceeds exactly as in the deduction of (Eq. (5.129), replacing derivatives by differences, etc. If the scattering system is initially in the m th energy state, we have

$$\left.\frac{d\sigma}{d\Omega}\right]_{classical} \leftrightarrow \frac{1}{\hbar^2 c^4}\sum_n (\omega-\omega_{nm})^4 \Big/ \Big| \sum_r \frac{n\times\eta_{nr}\, l\cdot\eta_{rm}}{\omega-\omega_{rm}} - \frac{n\times\eta_{n-(r-n)n}\, l\cdot\eta_{n\,n-(r-n)}}{\omega-\omega_{n\,n-(r-n)}}$$

$$- n\times\eta_{m\,m-r\,m}\left[\frac{l\cdot\eta_{m\,m+n-r}}{\omega-\omega_{n\,m+n-r}} - \frac{l\cdot\eta_{n\,2m-r}}{\omega-\omega_{m\,2m-r}}\right]\Big\}\Big|^2$$

$$= \frac{1}{\hbar^2 c^4}\sum_n (\omega-\omega_{nm})^4 \Big/ \Big| \sum_r \left\{\frac{n\times\eta_{nr}\, l\cdot\eta_{rm}}{\omega-\omega_{rm}} - \frac{l\cdot\eta_{nr}\, n\times\eta_{rm}}{\omega-\omega_{nr}}\right\}\Big|^2 \quad (5.185)$$

where the summation of each term is to be carried out over as many values of r as possible.

The individual terms in the sum (5.185) have the following physical significance. The incoming photon of angular frequency ω is absorbed by the system and a photon of angular frequency $\omega-\omega_{nm}$ is emitted, the system passing at the same time from the m th energy state to the n th energy state. Energy is evidently conserved in the process. The energy of the incoming photon is $\hbar\omega$, that of the emitted photon is $\hbar(\omega-\omega_{nm})$ and the energy left in the sysyem is $\hbar\omega_{nm}$.

Equation (5.185) does not quite give the quantum differential cross section as it stands. In the quantum theory we are interested in the scattering of individual particles rather than of the energy flux. We define the incoming and scattered particle fluxes by

$$\overline{S_p^{in}} = \frac{\overline{S^{in}}}{\hbar\omega} \quad , \quad \overline{S_p^{scat}} = \frac{\overline{S^{scat}}}{\hbar(\omega-\omega_{nm})} \quad (5.186)$$

and the differential cross section for scattering of photons by

$$\frac{d\sigma}{d\Omega} = \frac{r^2\,\overline{S_p^{scat}}}{\overline{S_p^{in}}} = \frac{r^2\,\overline{S^{scat}}}{\overline{S^{in}}}\,\frac{\hbar\omega}{\hbar(\omega-\omega_{nm})} \quad . \quad (5.187)$$

To obtain the correct quantum expression we must evidently multiply the individual terms of (5.185) by $\omega/\omega-\omega_{nm}$. Thus

$$\frac{d\sigma}{d\Omega} = \frac{1}{\hbar^2 c^4}\sum_n \omega(\omega-\omega_{nm})^3 \Big/ \Big| \sum_r \left\{\frac{n\times\eta_{nr}\, l\cdot\eta_{rm}}{\omega-\omega_{rm}} - \frac{l\cdot\eta_{nr}\, n\times\eta_{rm}}{\omega-\omega_{nr}}\right\}\Big|^2 \quad (5.188)$$

Equation (5.188) is known as the Kramers-Heisenberg dispersion formula. It accounts very well for the experimental data regarding the dispersion of light by atoms. It will be observed, both from the classical expression (5.184) and the quantum expression (5.188) that if the angular frequency of the incoming radiation happens to coincide with any of the natural frequencies of the scattering system then the scattering becomes very large owing to the vanishing of denominators. This situation is known as resonance fluorescence.

As a final illustration in this chapter of the application of the Correspondence Principle, we shall consider the scattering of electromagnetic radiation by a single charged particle. This example may be regarded as a limiting case of scattering by a bound multiply periodic system, which has been treated immediately above. As the binding force on the particle becomes weaker and weaker, the particle may in the limit be regarded as free. It will be observed that this will be the first example; to be treated quantitatively by us, of a quantized free system. S_o far the quantization rules have been defined only for multiply periodic systems. It will be one of our tasks in subsequent chapters to extend the quantum theory to non-periodic, or "free", systems. The present example, however, will show us that we already have enough mathematical equipment to gain a partial insight into what the nature of the quantum theory of non-periodic systems will be like..

The problem of radiation scattering by a single free charged particle has been treated classically in chapter 3. There we obtained expressions for the mean energy and momentum densities of the incoming plane monochromatic wave. In the quantum theory these densities must be related to probable photon density P in the following manner :

$$P = \frac{\overline{U^{in}}}{\hbar\omega} = \frac{\overline{|G^{in}|}}{\hbar|k|} \qquad (5.189)$$

Here $\hbar k$ is the momentum of an individual photon, k being a quantity

which we have not heretofore introduced. Since, from (3.171), (3.176), (3.182),

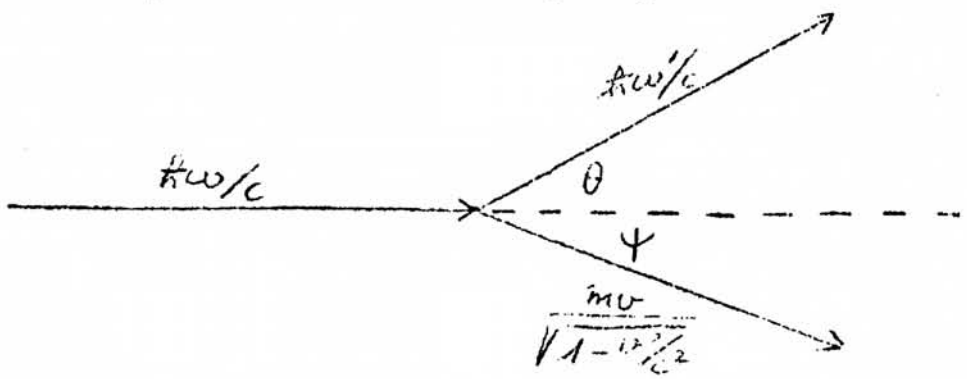
$$\frac{|\overline{G^{th}}|}{U^{in}} = \frac{1}{c}, \tag{5.190}$$

we evidently have

$$|k| = \frac{\omega}{c} = \frac{2\pi}{\lambda} \tag{5.191}$$

where λ is the wavelength of the photon.

In the quantum theory the scattering process is to be pictured as the absorption of an incident photon of momentum $\hbar\omega/c$ by a charged particle which is initially at rest, together with the emission of another photon, of momentum $\hbar\omega'/c$, in a different direction. Since experiment shows that momentum, as well as energy, is conserved in the process, the scattering particle must suffer a recoil and be found moving off in its own direction with a certain velocity v after the interaction with the photon. The scattering process is pictured in the following diagram.



The relations between the various angles and magnitudes may be readily calculated. From conservation of energy we get

$$\hbar\omega + mc^2 = \hbar\omega' + \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{5.192}$$

and from conservation of momentum we get

$$\frac{\hbar\omega}{c} = \frac{\hbar\omega'}{c}\cos\theta + \frac{mv}{\sqrt{1-v^2/c^2}}\cos\psi, \tag{5.193}$$

$$0 = \frac{\hbar\omega'}{c}\sin\theta - \frac{mv}{\sqrt{1-v^2/c^2}}\sin\psi. \tag{5.194}$$

We manipulate these equations as follows.

$$m^2c^2 + 2mc\,\frac{\hbar}{c}(\omega-\omega') + \frac{\hbar^2}{c^2}(\omega-\omega')^2 = \frac{m^2c^2}{1-\frac{v^2}{c^2}} \tag{5.195}$$

$$\frac{\hbar^2}{c^2}(\omega-\omega'\cos\theta)^2 = \frac{m^2v^2}{1-\frac{v^2}{c^2}}(1-\sin^2\psi)$$

$$= \frac{m^2v^2}{1-v^2/c^2} - \frac{\hbar^2}{c^2}\omega'^2\sin^2\theta \tag{5.196}$$

or

$$\frac{m^2v^2}{1-\frac{v^2}{c^2}} = \frac{\hbar^2}{c^2}(\omega^2+\omega'^2-2\omega\omega'\cos\theta). \tag{5.197}$$

Hence

$$\frac{m^2c^2}{1-v^2/c^2}\left(1-\frac{v^2}{c^2}\right) = m^2c^2 + 2m\hbar(\omega-\omega') + \frac{\hbar^2}{c^2}\left[(\omega-\omega')^2-\omega^2-\omega'^2+2\omega\omega'\cos\theta\right] \tag{5.198}$$

or

$$m\hbar(\omega-\omega') = \frac{\hbar^2}{c^2}\omega\omega'(1-\cos\theta) \tag{5.199}$$

Dividing by $m\hbar\omega\omega'$, we obtain

$$\frac{1}{\omega'} - \frac{1}{\omega} = \frac{\hbar}{mc^2}(1-\cos\theta) \tag{5.200}$$

Multiplying by $2\pi c$, we may reexpress this equation in terms of

the wavelengths of the incident and scattered photons.

$$\lambda' - \lambda = 2\pi \lambda\!\!\!^{-} (1 - \cos\theta) \tag{5.201}$$

where

$$\lambda\!\!\!^{-} = \frac{\hbar}{mc} \tag{5.202}$$

$\lambda\!\!\!^{-}$ is called the Compton wavelength of the scattering particle, the whole quantum process, whereby the scattering particle suffers a recoil and the photon suffers a change in wavelength as a result of their interaction, being known as the Compton effect. When the photon is scattered through 90° the wavelength shift is equal to $2\pi\lambda\!\!\!^{-}$. This shift has been very carefully measured in the laboratory.

It is of interest to make note at this point of a rather interesting relationship which exists between the three universal lengths associated with an electron, namely the classical electron radius r_0, the Bohr radius r_B, and the Compton wavelength $\lambda\!\!\!^{-}$. From equations (3.132), (5.36), (5.40) and (5.202), we get

$$r_B : \lambda\!\!\!^{-} : r_0 = 1 : \alpha : \alpha^2 \tag{5.203}$$

where α is the fine structure constant.

The wavelength shift expressed by eq. (5.201) may be interpreted as a Doppler shift. We have seen classically that the scattering particle is made to oscillate by the incoming electromagnetic field and that it also gradually picks up a drift velocity $\frac{\alpha^2}{2}c$ in the direction of propagation of the incoming field, the acceleration in this direction being given by eq. (3.186). This drift velocity produces a Doppler shift in the wavelength of the field radiated by the scattering particle. The shift depends on the angle θ at which the scattered radiation is observed, and can be determined by the

following reasoning. First, the particle is receding from the incoming radiation and hence it experiences a field of angular frequency lower than that of the incoming field by an amount

$$\Delta\omega = \frac{1}{c}\frac{d\bar{z}}{dt}\omega \qquad \left(\frac{d\bar{z}}{dt} \ll c\right) \tag{5.204}$$

Then, in the process of radiation by the particle, another shift comes in, which cancels the first one in the forward direction but doubles it in the backward direction. The angular frequency of the scattered radiation in the direction θ is

$$\omega' = (\omega - \Delta\omega)\left(1 + \frac{1}{c}\frac{d\bar{z}}{dt}\cos\theta\right)$$

$$= \omega - \frac{\omega'}{c}\frac{d\bar{z}}{dt}(1-\cos\theta) \qquad \frac{d\bar{z}}{dt} \ll c \tag{5.205}$$

Multiplying by $\frac{2\pi c}{\omega\omega'}$, we obtain

$$\lambda' - \lambda = \lambda\frac{1}{c}\frac{d\bar{z}}{dt}(1-\cos\theta) \tag{5.206}$$

Eq. (5.206) has exactly the same form as (5.201). From the Correspondence Principle we should expect that the value of $\frac{d\bar{z}}{dt}$ for which eqs. (5.201) and (5.206) become identical should be somewhere in between zero and the value which $\frac{d\bar{z}}{dt}$ has when one whole quantum has been scattered out of the incoming field. Actually, it turns out that the appropriate value of $\frac{d\bar{z}}{dt}$ is exactly that for which one quantum has been scattered. To show this we refer to eqs. (3.179), (3.180) and (3.186). The length of time required to scatter one quantum is

$$\Delta t = \frac{\hbar\omega}{c\,\bar{S}_3^{in}} = \frac{\hbar\omega}{\frac{8\pi}{3}\frac{e^4}{m^2c^4}\frac{1}{1+\omega^2\tau^2}\frac{1}{4\pi}\frac{\omega^2}{c^2}(\delta A)^2}$$

$$= \frac{\hbar\omega}{mc\,a_s} \tag{5.207}$$

The drift velocity of the scattering particle after this length of

time has passed is

$$\frac{d\bar{E}}{dt} = a_s \Delta t = \frac{\hbar\omega}{mc} = \frac{2\pi}{\lambda}\frac{\hbar}{m} \qquad (5.208)$$

Substitution of (5.208) into (5.206) gives (5.201).

The condition for the validity of the scattering formulae obtained in chapter 3, namely $\frac{d\bar{E}}{dt} \ll c$, imposes a restriction on the values of the angular frequency ω for which these formulae are valid. From (5.208) this restriction is immediately seen to be given by

$$\hbar\omega \ll mc^2 \qquad (5.209)$$

That is, the energy of the incident photon must be much less than the rest energy of the scattering particle. When its energy is comparableto or greater than the particle rest energy we may expect considerable deviations from the cross sections calculated in chapter 3. The Compton formula (5.201), however, will always be valid, since it is derived on a rigorous relativistic basis, independently of any assumptions about the form of the differential cross section.

Bibliography

M. Born, The Mechanics of the Atom, London, G. Bell and Sons, Ltd. 1927

W. Heitler, The Quantum Theory of Radiation, Oxford, 1947

D. Bohm, Quantum Theory, New-York, Prentice-Hall, Inc. , 1951

6.- HEISENBERG THEORY AND THE INTERPRETATION.

In chapter 5 we introduced certain quantities η_{nm} as quantum analogs of the Fourier amplitudes η_τ of the electric dipole moment of a charged multiply periodic system. By applying the Correspondence Principle we showed that it was possible to make rough estimates of the values of these quantities. We now wish to develop an exact quantum theory which will enable us to assign precise values to these quantities. More generally, we shall wish to obtain the quantum analogs of the Fourier amplitudes of any classically observable quantity.

Let G be any classical observable which is a single valued function of the canonical variables q^i, p_i defined with respect to some Lagrangian function. If the system is multiply periodic then F may be expanded in terms of the angle and action variables

$$F = \sum_\tau F_\tau(J) e^{i\tau\omega} \tag{6.1}$$

It is seen that the classical quantity F is completely determined by specifying the elements of an ordered ensemble, namely

$$(F(\tau, J)) \tag{6.2}$$

where

$$F(\tau, J) \equiv F_\tau(J) e^{i\tau\omega(J)t} \tag{6.3}$$

If the phases of the angle variables are chosen so that

$$\omega = \omega t \tag{6.4}$$

from which it is immediately evident that

$$(FG)(\tau,\gamma) = \sum_{\tau'} F(\tau-\tau',\gamma)\, G(\tau',\gamma)$$
$$= \sum F_{\tau-\tau'} G_{\tau'} e^{i(\tau-\tau')\omega t} e^{i\tau'\omega t}. \qquad (6.8)$$

As a quantum analog of (6.8) we may take

$$\langle E_n | FG | E_m \rangle = \sum_r \langle E_n | F | E_r \rangle \langle E_r | G | E_m \rangle$$
$$= \sum_r F_{nr} G_{rm} e^{i\omega_{nr} t} e^{i\omega_{rm} t}$$
$$= \sum_r F_{nr} G_{rm} e^{i\omega_{nm} t}. \qquad (6.9)$$

The multiplication law for ensembles of the type (6.5) is then seen to be that of ordinary matrix multiplication. In fact, the quantum ensembles may be regarded in every way simply as matrices. It is for this reason that quantum mechanics was often referred to in the early days as "matrix mechanics". It is easy to show that the quantum ensembles must also obey the other algebraic rules of matrices. Thus, since for the classical ensembles we have

$$(F+G)(\tau,\gamma) = F(\tau,\gamma) + G(\tau,\gamma) \qquad (6.10)$$

and

$$(aF)(\tau,\gamma) = aF(\tau,\gamma), \qquad (6.11)$$

where a is an arbitrary complex number, we must have, for the quantum ensembles,

$$\langle E_n | F+G | E_m \rangle = \langle E_n | F | E_m \rangle + \langle E_n | G | E_m \rangle \quad (6.12)$$

and

$$\langle E_n | aF | E_m \rangle = a \langle E_n | F | E_m \rangle \quad (6.13)$$

It should now be observed that there is an ambiguity in passing from the classical multiplication law (6.8) via the Correspondence Principle to the quantum multiplication law (6.9). We could equally well have written (6.9) in the form

$$\langle E_n | GF | E_m \rangle = \sum_\mu \langle E_n | G | E_\mu \rangle \langle E_\mu | F | E_m \rangle \quad (6.14)$$

and expressions (6.9) and (6.14) are not in general equal. Matrices obey a non-commutative law of multiplication, and we see that in the quantum theory it will be necessary to make a distinction between the products FG and GF. Just exactly what the quantum analog of a classical product FG should be, whether FG, GF, some combination of the two, or even something else, therefore becomes problematical. We shall see in a subsequent chapter how this ambiguity can be removed in all cases which are of interest in physics, but for the present we shall sidestep this question. Rather, we shall study the general non-commutativity of quantum factors in a different light.

Having introduced matrix algebra into the quantum theory we must next seek to develop a dynamical theory of the quantum matrices as a means of eventually determining the precise values of the matrix elements. Again the Correspondence Principle guides us. We must parallel as closely as possible classical dynamical theory. Among the most important conceptual tools used in classical dynamical theory are the canonical invariants; and among the most important of these are the Poisson brackets. We should expect analogous invariants to play an important role in the quantum theory. Therefore, let us try to find a quantum analog of the classical Poisson brackets.

In terms of the angle and action variables the Poisson bracket of two classical observables, F and G, is given by

$$(F,G) = \sum_{\tau,\tau'} (F_\tau e^{i\tau\omega}, G_{\tau'} e^{i\tau'\omega})$$

$$= \sum_{\tau,\tau'} -i\tau_i \frac{\partial F_\tau}{\partial J_i} G_{\tau'} + i\tau_i F_\tau \frac{\partial G_{\tau'}}{\partial J_i} \Big) e^{i(\tau+\tau')\omega} \qquad (6.15)$$

The elements of the ordered ensemble corresponding to this Poisson bracket are evidently given by

$$(F,G)(\tau,\gamma) = \sum_{\tau'} -i\tau_i \frac{\partial F_{\tau-\tau'}}{\partial J_i} G_{\tau'} + i(\tau_i - \tau_i') F_{\tau-\tau'} \frac{\partial G_{\tau'}}{\partial J_i} \Big) e^{i\tau\omega t}$$

$$= \sum_{\tau'} \Big(-i\tau_i \frac{\partial}{\partial J_i} (F_{\tau-\tau'} G_{\tau'}) + i\tau_i F_{\tau-\tau'} \frac{\partial G_{\tau'}}{\partial J_i} \Big) e^{i\tau\omega t} \qquad (6.16)$$

The quantum analogue of this equation may be obtained in a manner closely similar to that of the derivation of the Kramers-Heisenberg dispersion formula (5.185). We have

$$\langle E_n | (F,G) | E_m \rangle$$

$$= -\frac{i}{\hbar} \sum_r \Big\{ F_{nr} G_{rm} - F_{n-(r-n)\,n} G_{m-(r-m)\,m} - F_{m+n-r\,n} \Big(G_{m\,m+n-r} - G_{m\,n-(n-r)} \Big) \Big\} e^{i\omega_{nm} t}$$

$$= \frac{1}{i\hbar} \sum_r \Big\{ F_{nr} e^{i\omega_{nr} t} G_{rm} e^{i\omega_{rm} t} - G_{nr} e^{i\omega_{nr} t} F_{rm} e^{i\omega_{rm} t} \Big\}$$

$$= \frac{1}{i\hbar} \sum_r \Big\{ \langle E_n | F | E_r \rangle \langle E_r | G | E_m \rangle - \langle E_n | G | E_r \rangle \langle E_r | F | E_m \rangle \Big\} \qquad (6.17)$$

$$= \frac{1}{i\hbar} \langle E_n | [F,G] | E_m \rangle$$

where $[F, G]$ denotes the "commutator" of F and G.

$$[F,G] = FG - GF. \tag{6.18}$$

We see, therefore, that the commutator bracket in quantum mechanics will play the same role as the Poisson bracket in classical mechanics. Explicitly we have

$$(F,G) = \frac{1}{i\hbar}[F,G] \tag{6.19}$$

In writing equations of the form (6.18, 19) we may think of the symbols F and G as standing for the matrices themselves of which the quantities (6.6) etc, are the elements.

The most logical way of developing a dynamical theory of the quantum matrices is to postulate that their time behavior is given by the usual canonical equation, now, however, transcribed into quantum form, namely

$$\dot{F} = (F,H) = \frac{1}{i\hbar}[F,H], \tag{6.20}$$

where H is the matrix corresponding to the Hamiltonian of the system in question. The postulate actually consists in the statement that the quantum matrix corresponding to the time derivative of a classical quantity F is identical with the time derivative of the matrix corresponding to F itself. That is,

The time derivative of the matrix = The matrix of the time derivative.

For if $\dot{F}$ is understood to stand for the time derivative of the matrix, then the validity of equation (6.20) is already contained in the formalism of chapter 5. To show this we must first find the matrix corresponding to the Hamiltonian function.

In the classical theory

$$H = E(\vec{J}) \tag{6.21}$$

That is, in the Fourier expansion of the Hamiltonian function in terms of the angle and action variables, only the constant term appears. There are no oscillating terms. This means

$$H_{nm} = 0 \qquad \text{for } m \neq n \tag{6.22}$$

Moreover, by (5.45) we have

$$H_{nm} = E(J_n) = E_n \qquad N.S. \tag{6.23}$$

Hence

$$\langle E_n | H | E_m \rangle = \delta_{nm} E_n = \delta_{nm} E_m \qquad N.S. \tag{6.24}$$

If we now write the right side of (6.20) in explicit matrix form, we obtain

$$\begin{aligned}
&\frac{1}{i\hbar} \langle E_n | [F, H] | E_m \rangle \\
&= \frac{1}{i\hbar} \sum_r \left\{ \langle E_n | F | E_r \rangle \langle E_r | H | E_m \rangle - \langle E_n | H | E_r \rangle \langle E_r | F | E_m \rangle \right\} \\
&= \frac{1}{i\hbar} \sum_r \left\{ E_m \delta_{rm} F_{nr} e^{i\omega_{nr} t} - E_r \delta_{nr} F_{rm} e^{i\omega_{rm} t} \right\} \\
&= \frac{E_n - E_m}{i\hbar} F_{nm} e^{i\omega_{nm} t} = i\omega_{nm} F_{nm} e^{i\omega_{nm} t} \\
&= \frac{d}{dt} \langle E_n | F | E_m \rangle \qquad \text{q.e.d.}
\end{aligned} \tag{6.25}$$

Of particular interest are the commutation relations for the matrices corresponding to the canonical coordinates and momenta. We have

$$[\rho^i, \rho^j] = 0, \quad [p_i, p_j] = 0, \quad [\rho^i, p_j] = i\hbar\delta^i_j \tag{6.26}$$

These commutation relations may be used to demonstrate the validity of an interesting application of the Correspondence Principle to the Kramers-Heisenberg dispersion formula (5.188). Suppose that the multiply periodic system to which the Kramers-Heisenberg dispersion formula is applied consists simply of a single charged particle bound in some kind of static potential field. Then the dipole moment of the system is given by

$$\eta = e\rho \tag{6.27}$$

where e is the charge on the particle and ρ is the radius vector to the particle from some origin, say the equilibrium position. If the components of ρ are taken as the canonical coordinates then the conjugate momenta are the compoonents of the momentum vector p, where

$$p = m\dot{\rho}, \tag{6.28}$$

m being the mass of the particle. In quantum matrix form equations (6.27, 28) become

$$\eta_{n,m} = e\rho_{n,m} \tag{6.29}$$

$$p_{n,m} = im\omega_{n,m}\rho_{n,m}. \tag{6.30}$$

Using eqs (6.6), (6.26), (6.29), (6.30) we find

$$\sum_r (\eta_{inr}\eta_{jrm} - \eta_{jnr}\eta_{irm})$$
$$= \sum_r e^2 \{\langle E_n|r_i|E_r\rangle\langle E_r|r_j|E_m\rangle - \langle E_n|r_j|E_r\rangle\langle E_r|r_i|E_m\rangle\} e^{-i\omega_{nm}t} = 0 \quad (6.31)$$

and

$$\sum_r (-\omega_{nr}\eta_{inr}\eta_{jrm} + \omega_{rm}\eta_{jnr}\eta_{irm})$$
$$= \sum_r \left(-\frac{e^2}{2im} p_{inr} r_{jrm} + \frac{e^2}{im} r_{jnr} p_{irm}\right)$$
$$= \frac{e^2}{im}\sum_r \{\langle E_n|r_j|E_r\rangle\langle E_r|p_i|E_m\rangle - \langle E_n|p_i|E_r\rangle\langle E_r|r_j|E_m\rangle\} e^{-i\omega_{nm}t}$$
$$= \frac{e^2\hbar}{m}\delta_{ij}\delta_{nm} e^{-i\omega_{nm}t} = \frac{e^2\hbar}{m}\delta_{ij}\delta_{nm}. \quad (6.32)$$

Now, classically, as the angular frequency ω of the incoming radiation becomes higher and higher the effect of binding on the charged particle becomes less and less important, until, in the limit in which ω becomes much larger than any of the natural frequencies of the particle, the particle may be regarded as essentially free. The Correspondence Principle tells us that the same should be true in the quantum theory. That is, in the limit of large ω the Kramers-Heisenberg dispersion formula should reduce to the formula which gives the differential cross section for scattering of light by a free particle. With the use of eqs. (6.31, 32) we may show that this is, in fact, the case.

In the limit as ω becomes much larger than any of the ω_{nr}, equation (5.188) becomes

$$\frac{d\sigma}{d\Omega} \to \frac{1}{\hbar^2 c^4}\sum_{n,i,j,k} \omega^4 |\varepsilon_{ij}|^2 n_i \rho_k \sum_r \left\{\frac{(\omega-\omega_{rn})\eta_{jnr}\eta_{krm} - (\omega-\omega_{rm})\eta_{jnr}\eta_{krm}}{\omega^2}\right\}$$

$$= \frac{1}{\hbar^2 c^4} \sum_{n,i,j,k} \left| \varepsilon_{ij} n_i l_k \sum_{s} \left\{ \omega \left(\eta_{jsn} \eta_{ksm} - \eta_{ksn} \eta_{jsm} \right) + \left(\omega_{sn} \eta_{jsn} \eta_{ksm} + \omega_{sm} \eta_{ksn} \eta_{jsm} \right) \right\} \right|^2$$

$$= \frac{e^2}{mc^4} \sum_{n,i,j,k} \left| \varepsilon_{ij} n_i l_k \delta_{jk} \delta_{nm} \right|^2 = \frac{e^4}{m^2 c^4} (n \times l)^2 \qquad (6.33)$$

Since $|n \times l| = \sin \beta$, where β is the angle between the scattering direction and the polarization direction, one sees, on comparing (6.33) with the result of Problem XXV, that the Kramers-Heisenberg cross section indeed reduces to that for the scattering of light by a free particle, with, however, neglect of radiation damping.*

If F in the classical theory is a real quantity then

$$F_\tau^* = F_{-\tau} \qquad (6.34)$$

The quantum analogue of this reality condition is (cf. (5.46)

$$(F_{nm})^* = F_{mn} \qquad (6.35)$$

or, in full matrix form, since $\omega_{nm} = -\omega_{mn}$,

$$\langle E_n | F | E_m \rangle^* = \langle E_m | F | E_n \rangle . \qquad (6.36)$$

---

* The effect of particle recoil and the resultant Doppler shift $\omega \to \omega'$ is also neglected. To obtain the correct quantum cross section one should multiply expression (6.33) by ω / ω' according to the same reasoning used in passing from Eq. (5.185) to Eq. (5.188). The present scheme, in fact, cannot take the Doppler shift into account as this shift depends on the translational motion of the free particle. This translational motion can be taken into account only if the initial state is a pure momentum state, which, in this bound state analysis, it is not.

That is, the quantum matrix corresponding to F is Hermitian. The reality condition on an observable will be expressed in both the classical and quantum theories in the form

$$F^* = F. \tag{6.37}$$

It is of interest to examine the form which classical canonical transformations take in the quantum theory. We first consider infinitesimal transformations. If εS is the infinitesimal generating function of an infinitesimal canonical transformation then the change ΔF in an observable F under this transformation is given by the equation (1.56). The quantum analog of this equation is

$$\Delta F = \frac{\varepsilon}{i\hbar}[F, S] \tag{6.38}$$

To the first infinitesimal order we may write

$$F + \Delta F = \left(1 - \frac{\varepsilon}{i\hbar} S\right) F \left(1 + \frac{\varepsilon}{i\hbar} S\right) = U^{-1} F U \tag{6.39}$$

where

$$U = 1 + \frac{\varepsilon}{i\hbar} S. \tag{6.40}$$

Since S is real (i.e. Hermitian), U is an infinitesimal unitary matrix and equation (6.39) expresses an infinitesimal unitary transformation. Since the product of two unitary transformations is also a unitary transformation, it is evident that a finite canonical transformation generated by infinitesimal canonical transformations corresponds, in the quantum theory, to a finite unitary transformation generated by infinitesimal unitary transformations. In general, therefore, we see that a canonical transformation in classical theory corresponds to a unitary transformation in quantum theory.

The matrices corresponding to physical observables will usually be infinite Hermitian matrices. In the theory of finite matrices it is well known that any Hermitian matrix can be transformed to diagonal form by some unitary matrix. It will be assumed that this is also always true for the Hermitian matrices of quantum mechanics even though they may be infinite. It will often be of interest to find a unitary matrix which will transform a given Hermitian quantum matrix to diagonal form. An excellent illustrative example of a typical procedure for doing this is afforded by the derivation of the quantum analog of the classical perturbation theory developed in chapter 2.

First observe that the energy matrix (6.24) is already diagonal. If the angle and action variables of the classical system are known, we have seen in chapter 5 how we can determine (at least approximately) the non-vanishing diagonal elements of this energy matrix. If, however, the angle and action variables are not exactly known, as is the case in a problem to which perturbation methods must be applied, and if the Hamiltonian function is expressed in terms of a set of canonical variables which are only approximately equal to the true angle and action variables, then the quantum energy matrix constructed on the basis of these incorrect variables will not be diagonal. Thus, the quantum matrix form of equation (2.1) is

$$\langle E_{on}|H|E_{om}\rangle = \langle E_{on}|H_0|E_{om}\rangle + \langle E_{on}|H_1|E_{om}\rangle \quad (6.41)$$
$$= E_{om}\,\delta_{nm} + \langle E_{on}|H_1|E_{om}\rangle .$$

and since the perturbation function H_1 is not, in general, independent of the variables ω_0^i, the matrix $(\langle E_{on}|H_1|E_{om}\rangle)$ will have non-vanishing off-diagonal elements. The E_{on} are here the energy levels of the unperturbed system.

The procedure, in the classical theory, of finding a canonical transformation leading to the true angle and action variables will, in the quantum theory, become a procedure of finding a unitary transformation which diagonalizes the energy matrix. As in chapter 2,

we shall use a method of successive approximations. We shall, in fact, parallel as closely as possible the procedures of chapter 2.

When the classical Hamiltonian function H is expressed in terms of the true angle and action variables α^i, β_i instead of the approximate angle and action variables α_0^i, β_{0i}, its functional form changes. The amount of this change is, to first order, given by

$$\Delta H = -(H, W_1), \tag{6.42}$$

(see equations (1.60), (2.32) and (2.34)). If we denote by H' the altered function, then we may write

$$H' = H - (H, W_1) + \dots \tag{6.43}$$

The quantum analog of this equation is

$$\begin{aligned} H' &= H - \frac{1}{i\hbar}[H, W_1] + \dots \\ &= U^{-1} H U \end{aligned} \tag{6.44}$$

where

$$U = 1 - \frac{1}{i\hbar} W_1 + \dots \tag{6.45}$$

In the classical theory H' is independent of the angle variables. In the quantum theory the matrix H' is diagonal. U is therefore the unitary matrix which diagonalizes H.

To the first order of approximation we may write

$$H' = H_0 + H_1 - \frac{1}{i\hbar}[H_0, W_1] + \dots \tag{6.46}$$

or, in full matrix form,

$$\langle E_n|H|E_m\rangle = \langle E_{on}|H_0|E_{om}\rangle + \langle E_{on}|H_1|E_{om}\rangle - \frac{1}{i\hbar}\sum_r \{\langle E_{on}|H_0|E_{or}\rangle\langle E_{or}|W_1|E_{om}\rangle - \langle E_{on}|W_1|E_{or}\rangle\langle E_{or}|H_0|E_{om}\rangle\} + \ldots \quad (6.47)$$

which reduces to

$$E_m\delta_{nm} = E_{om}\delta_{nm} + \langle E_{on}|H_1|E_{om}\rangle - \frac{E_{on}-E_{om}}{i\hbar}\langle E_{on}|W_1|E_{om}\rangle + \ldots \quad (6.48)$$

In order that the expressions on both sides of this equation involve the Kronecher delta $\delta_{n,m}$ in the same way we must have

$$\langle E_{on}|W_1|E_{om}\rangle = i\hbar\frac{\langle E_{on}|H_1|E_{om}\rangle}{E_{on}-E_{om}} = i\frac{\langle E_{on}|H_1|E_{om}\rangle}{\omega_{onm}} \quad \text{for } m\neq n \quad (6.49)$$

or, dropping the exponential factors $e^{i\omega_{onm}t}$ which accompany all matrix elements,

$$W_{1nm} = \frac{iH_{1nm}}{\omega_{onm}} \quad \text{for } m\neq n \quad (6.50)$$

Equation (6.50) is clearly the quantum analog of (2.25). We may, without loss of generality, take

$$W_{1mm} = 0 \quad \text{for all } m \quad \text{N.S.} \quad (6.51)$$

which is the quantum analog of (2.14).

If we write (cf. (2.6))

$$E_m = E_{om} + E_{1m} + E_{2m} + \ldots \quad (6.52)$$

we have, on setting m = n in (6.48),

$$E_{1m} = \langle E_{0m} | H_1 | E_{0m} \rangle = H_{1mm} \quad N.S. \tag{6.53}$$

which is the quantum analog of (2.22).

The second order corrections to the energy levels of the quantized system may also be readily calculated. For this purpose we need to evaluate the matrix U correct to the second order. The requirement that U be unitary turns out to make this evaluation unique. We must have, in fact

$$U = 1 - \frac{1}{i\hbar}(W_1 + W_2) - \frac{1}{2\hbar^2} W_1^2 + \ldots \tag{6.54}$$

$$U^{-1} = 1 + \frac{1}{i\hbar}(W_1 + W_2) - \frac{1}{2\hbar^2} W_1^2 + \ldots \tag{6.55}$$

Here W_2 is the quantum analog of the corresponding classical quantity. Its explicit evaluation will not actually be needed in the calculation of E_{2m}. Equation (6.44) now becomes, correct to second order,

$$H' = H_0 + H_1 - \frac{1}{i\hbar}[H_0 + H_1, W_1] - \frac{1}{i\hbar}[H_0, W_2] + \frac{1}{\hbar^2} W_1 H_0 W_1 - \frac{1}{2\hbar^2}\{H_0, W_1^2\} + \ldots \tag{6.56}$$

where the brackets $\{\ \}$ are used to denote the "anticommutator". Writing equation (6.56) in full matrix form and using (6.49) and (6.53), we have

$$E_{2m}\delta_{nm} = -\frac{1}{\hbar}\sum_{r \neq m,n}\left\{\langle E_{0n}|H_1|E_{0r}\rangle \frac{\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0rm}} - \frac{\langle E_{0n}|H_1|E_{0r}\rangle\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0nr}}\right\}$$

$$- \frac{E_{0n} - E_{0m}}{i\hbar}\langle E_{0n}|W_2|E_{0m}\rangle$$

$$- \frac{1}{\hbar^2}\sum_{r \neq m,n} \frac{\langle E_{0n}|H_1|E_{0r}\rangle}{\omega_{0nr}} E_{0r} \frac{\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0rm}}$$

$$+\frac{1}{2\hbar^2}\sum_{\substack{r\\ r\neq m,n}}\left(E_{0n}+E_{0m}\right)\frac{\langle E_{0n}|H_1|E_{0r}\rangle\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0nr}\,\omega_{0rm}} \tag{6.57}$$

Setting m = n in (6.57) we finally get

$$E_{2m}=-\frac{2}{\hbar}\sum_{\substack{r\\ r\neq m}}\frac{\langle E_{0m}|H_1|E_{0r}\rangle\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0rm}}$$

$$+\frac{1}{\hbar^2}\sum_{\substack{r\\ r\neq m}}\frac{E_{0r}-E_{0m}}{\hbar}\,\frac{\langle E_{0m}|H_1|E_{0r}\rangle\langle E_{0r}|H_1|E_{0m}\rangle}{\omega_{0rm}^2}$$

$$=-\frac{1}{\hbar}\sum_{r}\frac{|\langle E_{0r}|H_1|E_{0m}\rangle|^2}{\omega_{0rm}}$$

$$=-\frac{1}{\hbar}\sum_{r\neq m}\frac{|H_{1\,r-m}|^2}{\omega_{0rm}} \tag{6.58}$$

Equation (6.58) may easily be seen to be the quantum analog of the classical equation (2.26). We have

$$-\sum_{\tau\omega_0>0}\tau_i\frac{\partial}{\partial J_i}\left(\frac{|H_{1\tau}|^2}{\tau\omega_0}\right)$$

$$\rightarrow-\frac{1}{\hbar}\sum_{\substack{r\\ \omega_{rm}>0}}\left\{\frac{|H_{1\,r-m}|^2}{\omega_{0rm}}-\frac{|H_{1\,m\;m-(r-m)}|^2}{\omega_{0m\;m-(r-m)}}\right\}$$

$$=-\frac{1}{\hbar}\sum_{\substack{r\\ \omega_{rm}>0}}\left\{\frac{|H_{1\,r-m}|^2}{\omega_{0rm}}+\frac{|H_{1\,2m-r\,m}|^2}{\omega_{0\,2m-r\,m}}\right\} \tag{6.59}$$

which leads to (6.58).

So far we have used the Heisenberg matrices only as tools for calculating the probability of occurrence of certain physical processes (i.e. transitions), or for computing certain physical magnitudes such as energy levels. We have nowhere obtained a very clear pictu-

re as to their precise physical significance, other than to observe that they are somehow related to the Fourier amplitudes of the classical oscillations and this only when the energy matrix is diagonal. More precisely, we have nowhere obtained a very clear picture as to the precise physical nature of a quantum state.

Classically, a state of motion of a system is completely specified by giving the values of all the canonical variables (coordinates and momenta) at any instant of time. In a multiply periodic system this is equivalent to giving the values of all the action variables and the phases of all the angle variables. According to the old quantum theory a quantum energy state is specified by a particular combination of the possible discrete quantum values of the action variables. So far this mathematical specification of quantum states has sufficed us. A simple example, however, will show that this formulation is actually insufficient for a general quantum theory. Our job will be, on the basis of this example, to draw the outlines of a theory which will enable us to formulate the notion of "quantum state" with the same degree of mathematical generality as the notion of "state of motion" has been formulated in classical theory. In the course of this formulation the physical nature of a quantum state and the statistical basis of quantum theory will become a little clearer.

Consider a multiply periodic system which is described by a Hamiltonian function H_0. Let this system be perturbed, and let the perturbation be described by a function H_1, which now does not have to be regarded as small. Denote by H the Hamiltonian function of the perturbed system.

$$H = H_0 + H_1 \tag{6.60}$$

Suppose that an observation of the perturbed system tells us that the system is in the m th energy state. The system will remain

in this state until disturbed. Suppose, then, that at a subsequent time the perturbation is suddenly switched off and a measurement of the energy of the unperturbed system is simultaneously made. Classically, the motion of the unperturbed system subsequent to the "switch-off" will correspond to certain values of the action variables appropriate to that system. These values will not in general belong to the discrete set of permissable quantum values even though the action variables of the perturbed system were initially quantized. According to the quantum theory, however, the unperturbed system can be found only in a certain discrete energy state. How can we remove this apparent inconsistency from the theory and at the same time retain the Correspondence Principle between classical and quantum mechanics ?

Let us first observe that the values of the angle variables are not used in the specification of the quantum energy states of a system. This suggests, that in keeping with the statistical nature of quantum theory, we should average over the various values of the angle variables whenever they are not specified. This is equivalent to averaging over the time[*]. The expectation value of any quantity F in a state which is described only by giving values to the action variables J_i is therefore

$$\overline{F} = F_0 \tag{6.61}$$

---

[*] The fact that we use a time average here rather than some other average is because we are defining the quantum states in terms of energy values; since we are specifying the energy values exactly, the uncertainty relation (5.112) of chapter 5 tells us that we must have an infinite amount of time at our disposal in which to make this specification. The probable value of any observed quantity is, under these circumstances, its time average value.

where F_0 is the constant term in the classical Fourier expansion. The quantum analog of (6.61) is, of course,

$$\bar{F} = F_{mm} \tag{6.62}$$

In particular, the expectation value of the energy of the unperturbed system at "switch-off" is

$$\bar{H}_0 = H_{0mm} = \langle E_m | H_0 | E_m \rangle. \tag{6.63}$$

The actually observed value of H_0 will, however, in general differ from this, just as the classical value will generally differ from H_{00}, depending on the phases of the angle variables at measurement time. The quantum value will be equal to sometimes one and sometimes another of its allowed values. There will be a certain probability attached to each of the values, the mean or expectation value being given by (6.63).

Since the observed value of H_0 is not completely determinate, the unperturbed system cannot be said to be in a precise one of its quantum energy states at "switch-off" and measurement time. We must generalize the concept of "quantum state" so that when we speak of the quantum state of the unperturbed system we refer in some manner to a <u>mixture</u> of its possible energy states. In doing this we must be sure that the specification of a particular quantum state includes enough information to determine the probabilities of the energy states exactly.

It turns out that we are led to the necessary generalization through the notion of unitary transformation. Let U be the unitary matrix which diagonalizes the quantum matrix $(\langle E_m | H_0 | E_n \rangle)$ at the time of measurement:

$$\sum_{r,s} \langle E_{0m} | U^{-1} | E_s \rangle \langle E_s | H_0 | E_r \rangle \langle E_r | U | E_{0m'} \rangle = E_{0m} \delta_{mm'} \tag{6.64}$$

N.S.

Since U is unitary, we have

$$\langle E_{om}|U^{-1}|E_s\rangle = \langle E_s|U|E_{om}\rangle^* \tag{6.65}$$

$$\sum_r \langle E_r|U|E_{on}\rangle^* \langle E_r|U|E_{om}\rangle = \delta_{mn}, \tag{6.66}$$

$$\sum_r \langle E_n|U|E_{or}\rangle^* \langle E_m|U|E_{or}\rangle = \delta_{mn}. \tag{6.67}$$

Aside from possible phase factors the matrix U is uniquely determined by the labels E_m, E_{or}, etc. Hence it is customary to make the following abbreviations :

$$\langle E_r|U|E_{om}\rangle \equiv \langle E_r|E_{om}\rangle, \tag{6.68}$$

$$\langle E_{om}|U^{-1}|E_r\rangle \equiv \langle E_{om}|E_r\rangle, \tag{6.69}$$

the symbols U and U^{-1} being superfluons. From eqs. (6.64-6.67) we therefore have

$$\sum_{r,s} \langle E_{on}|E_s\rangle \langle E_s|H_0|E_r\rangle \langle E_r|E_{om}\rangle = E_{om}\delta_{mn}, \quad N.S., \tag{6.70}$$

$$\langle E_{on}|E_s\rangle = \langle E_s|E_{on}\rangle^* \tag{6.71}$$

$$\sum_r \langle E_r|E_{on}\rangle^* \langle E_r|E_{om}\rangle = \delta_{mn}, \tag{6.72}$$

$$\sum_r \langle E_n|E_{or}\rangle^* \langle E_m|E_{or}\rangle = \delta_{mn}. \tag{6.73}$$

The inverse of equation (6.70) is evidently

$$\langle E_n|H_0|E_m\rangle = \sum_r \langle E_n|E_{or}\rangle E_{or} \langle E_{or}|E_m\rangle \tag{6.74}$$

In particular, we have for the expectation value of H_0,

$$\langle E_m|H_0|E_m\rangle = \sum_r E_{or} |\langle E_{or}|E_m\rangle|^2. \tag{6.75}$$

The form of equation (6.75) strongly suggests that the quantity $|\langle E_{or}/E_{ms}\rangle|^2$ be interpreted as the probability that the value E_{om} will be obtained upon a measurement of the energy of the unperturbed system. The correctness of this interpretation can be shown in the following manner.

Let H_o^r be that function of H_o which takes on the value 1 when H_o is equal to E_{or} and the value 0 for all other values of H_o. The function H_o may for example be expressed in the formal manner

$$H_o^r = \delta_{H_o E_{or}} \tag{6.76}$$

Now, the matrix $(\langle E_r/E_{om}\rangle)$ which diagonalizes the quantum matrix H_o also simultaneously diagonalizes any function of H_o. In particular

$$\sum_{s,t} \langle E_{or}/E_t\rangle \langle E_t/H_o^r/E_s\rangle \langle E_s/E_{om}\rangle = \begin{cases} 1 & \text{for } m = n = r \\ 0 & \text{in all other cases} \end{cases} \tag{6.77}$$

It will next be observed that any measurement of H_o is also simultaneously a measurement of H_o^r. Owing to the special nature of H_o^r its expectation value will give directly the probability that a measurement of H_o will yield the value E_{or}. This is

$$\langle E_m/H_o^r/E_m\rangle = \langle E_m/E_{or}\rangle \langle E_{or}/E_m\rangle \qquad N.S.$$

$$= |\langle E_{or}/E_m\rangle|^2, \tag{6.78}$$

which yields the desired result.

It is evident that, as far as statistical considerations are concerned, the quantum state of the system after "switch-off" is completely determined by the quantities $\langle E_{or}/E_m\rangle$ with fixed m, that is, by the m th column of the matrix $(\langle E_{or}/E_m\rangle)$.

We shall now postulate that the quantum states of a system are determined in all respects by such columns of numbers. Such columns are commonly referred to as vectors -in particular, state vectors. For short we shall denote the m th column of the matrix $(\langle E_{or}/E_m\rangle)$ by $/E_m\rangle$. If the energy of the perturbed system is known to be E_m then the state of the system at all times is described by the vector $/E_m\rangle$.

If the observed values of the perturbed Hamiltonian function H are described in terms of the energy levels of the unperturbed system then one is led to an equation analogous to (6.70), namely

$$\sum_{r,s} \langle E_m/E_{os}\rangle \langle E_{os}/H/E_{or}\rangle \langle E_{or}/E_m\rangle = E_m \delta_{mn} \tag{6.79}$$

where $(\langle E_{os}/H/E_{or}\rangle)$ is the quantum matrix for H with respect to the unperturbed system. Equation (6.79) can also be written in the form

$$\sum_{r} \langle E_{os}/H/E_{or}\rangle \langle E_{or}/E_m\rangle = E_m \langle E_{os}/E_m\rangle. \tag{6.80}$$

That is, the vector $/E_m\rangle$ is the eigenvector of the matrix $(\langle E_{os}/H/E_{or}\rangle)$ corresponding to the eigenvalue E_m. It is now evident that the particular form of the Hamilto ian function H_o of the unperturbed system is unimportant. The equation (6.80) remains invariant in form no matter what the form of the function H_o is. This fact indicates that we should regard the quantum analog of a classical Hamiltonian function H as an operator which has an invariant significance independent of any particular matrix representation. A transformation from one representation to another can be made by means of a unitary matrix. Equation (6.80) can then be written in the symbolic form

$$H/E_m\rangle = E_m/E_m\rangle. \tag{6.81}$$

In a similar fashion we may write

$$H_o/E_{om}\rangle = E_{om}/E_{om}\rangle, \tag{6.82}$$

where $|E_{om}\rangle$ is the eigenvector of the quantum operator H_0 at measurement time corresponding to the eigenvalue E_{om}. Now it is evident, from a consideration of the case in which the perturbation function H_1 vanishes, that

$$\langle E_{on} | E_{om} \rangle = \delta_{mn}. \tag{6.83}$$

That is the unitary matrix reduces to the unit matrix. (Remember here that we are assuming the unperturbed system to be non-degenerate.) Hence

$$\sum_r \langle E_{on} | E_{or} \rangle \langle E_{or} | E_{rm} \rangle = \langle E_{on} | E_{rm} \rangle. \tag{6.84}$$

Equations (6.84) for the various values of n, may evidently be conbined into the vector form

$$\sum_r |E_{or}\rangle \langle E_{or} | E_{rm} \rangle = |E_{rm}\rangle. \tag{6.85}$$

Equation (6.85) gives an explicit representation of the state vector $|E_m\rangle$ as a linear combination of the eigenvectors $|E_{or}\rangle$ of the unperturbed Hamiltonian function H_0. The square of the absolute value of the coefficient $\langle E_{or} | E_{rm} \rangle$ has already been seen to be the probability that the unperturbed system will be found in the r th energy state if the perturbed system is known to be in the m th. Equation (6.85) expresses the actual state of the unperturbed system after "switch-off" as a <u>superposition</u> of its possible energy states, a certain weight, related to a physical probability, being attached to each energy state.

Energies are not the only quantities which can be measured, even in quantum systems. Therefore in order to extend the quantum theory to its fullest generality, we must associate an invariant quantum operator with every classical observable, and we must be able to talk about the eigenvectors of this operator. The eigen-

values of the operator will be the permissable quantum values of the corresponding observable.

With this extension of our original ideas it is possible to use the Heisenberg theory, directly as we have outlined it here, to derive practically all the results of quantum mechanics. However, it will be much more convenient if we first take the trouble to improve our notation (which has at times become rather cumbersome) and to develop some of the mathematics associated with the theory in a systematic way. In the following chapter it will be our task to do this.

Another one of our tasks will be to extend the quantum theory so as to apply it to systems which are not multiply periodic. Of course, any non-periodic system can be made periodic by placing it in a box. The larger the box the more nearly the enclosed system will approximate to the behavior of the free system. In principle therefore, the methods of the preceding chapters could be applied, together with a suitable limiting process, to non-periodic systems. It will turn out, however, that the development of the mathematical formalism in the next chapter will enable us to treat non-periodic systems in a much simpler way. The quantum treatment of degenerate systems will also become a straightforward process.

It should be emphasized finally, that in spite of the mathematical developments to follow, the introduction of the Heisenberg theory completes the full transition from classical mechanics - via the Correspondence Principle - to quantum mechanics.

Bibliography

M. Born and P. Jordan, Elementare Quantenmechanik, Berlin Julius Springer, 1930

Q U A N T U M M E C H A N I C S

Lecture Notes by
Bryce Seligman DEWITT
Radiation Laboratory
University of California

Chapter VII

Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

7. THE MATHEMATICS OF QUANTUM MECHANICS

One of the most convenient and widely used notations for denoting quantum mechanical quantities is that which has been developed by Dirac. We shall here give an outline of the Dirac formalism.

In the preceding chapter we have indicated that the theory of quantum phenomena can be formulated in an exact and rigorous manner by using complex vectors to describe the quantum states of physical systems. These vectors will be known as "state vectors". If we are not paying attention to the individual components of a state vector nor to any particular representation of the vector, but are only speaking of the abstract invariant property (or properties) which the vector describes, then an arbitrary state vector will be denoted by the symbol $|\ \rangle$. We have already hinted in the last chapter that such an abstract notation will be useful. One reason for this is the following. The number of dimensions of the space in which the state vectors exist will not necessarily be finite. For example, the number of energy levels of a quantized multiply periodic system is, in general, denumerably infinite, and hence the number of components of any state vector, in a representation in which the matrix of the Hamiltonian function is diagonal, is denumerably infinite. This means that the state-vector space has a denumerable infinity of dimensions. For other quantized systems, as we shall see later, the number of dimensions of the state-vector space is frequently even non-denumerably infinite. Therefore, if we attempted to deal with the individual components of a state vector in a particular representation, we should have to concern ourselves, among other things, with the questions of convergence. By making use of the abstract formalism, however, these questions can be by-passed until the time comes when it is necessary to consider them in detail for some particular system. It will be found, actually, that the physical properties of many systems can be mathematically deduced completely within the framework of the abstract formalism.

Following Dirac, we shall adopt a postulational approach based, physically, on the results of the preceding chapters and, mathematically, on analogy with the theory of ordinary finite-dimensional vector spaces. We have already introduced the bracket symbol $|\ \rangle$ to denote an arbitrary state vector. A particular state vector may be specified by inserting a label such as A inside the brackets, thus $|A\rangle$. If the space in which the state vectors exist is denoted by S, then its structure is characterized by the following set of postulates :

1) For any pair of vectors $|A\rangle$ and $|B\rangle$ in S there exists a vector in S called the sum of $|A\rangle$ and $|B\rangle$, which is denoted by the symbol $|A\rangle + |B\rangle$.

2) For any vector $|A\rangle$ in S and for any complex number a there exists a vector in S called the product of a and $|A\rangle$, which is denoted by the symbol $a\,|A\rangle$.

Sums and products obey the following rules :

3) $|A\rangle + |B\rangle = |B\rangle + |A\rangle$

4) $(|A\rangle + |B\rangle) + |C\rangle = |A\rangle + (|B\rangle + |C\rangle)$

5) $a(|A\rangle + |B\rangle) = a|A\rangle + a|B\rangle$

6) $(a+b)|A\rangle = a|A\rangle + b|A\rangle$

7) $a(b|A\rangle) = (ab)|A\rangle$

Since complex numbers commute, $ab = ba$, the vector quantity appearing in 7) is also equal to $b(a|A\rangle)$. Since the grouping of the symbols is evidently immaterial this vector will be denoted simply by $ab|A\rangle$. Similarly the vector quantity appearing in 4) will be denoted simply by $|A\rangle + |B\rangle + |C\rangle$.

8) There exists a null vector in S, denoted by 0, such that, for any $|A\rangle$, $|A\rangle + 0 = |A\rangle$, and $0\,|A\rangle = 0$. The symbol 0 appearing on the left side of the second equation is the number zero. No confusion will arise in practise

from using the same symbol to denote two different conceptual quantities.

9) $1|A\rangle = |A\rangle$.

A slight extension of the above notation is customarily made by introducing the abbreviations

$$-|A\rangle \equiv (-1)|A\rangle \tag{7.1}$$

$$|A\rangle - |B\rangle \equiv |A\rangle + (-|B\rangle) \tag{7.2}$$

Using these abbreviations, one may readily prove for example,

$$|A\rangle - |A\rangle = 0 \tag{7.3}$$

Problem XXXIV : Prove eq. (7.3)

The most characteristic property of the state vector space S is the provision which its structure makes for adding two state vectors together. This property corresponds to the notion of the superposition of quantum states which has already been suggested in chapter 6. To each state vector there corresponds a quantum state of the system in question. If any state vector $|T\rangle$ is expressible as a linear combination of other states $|A\rangle$, $|B\rangle$, etc.; in the form

$$|T\rangle = a|A\rangle + b|B\rangle + \ldots, \tag{7.4}$$

where a, b, etc. are complex numbers, then the quantum state corresponding to $|T\rangle$ will be said to be a superposition of the quantum states corresponding to $|A\rangle$, $|B\rangle$, etc.

We shall want the superposition of any quantum state with itself to result in the same state. The state vector

$$(a+b)|A\rangle = a|A\rangle + b|A\rangle \tag{7.5}$$

where a and b are arbitrary complex numbers, must therefore correspond to the same quantum state as does $|A\rangle$ itself. Since a + b is an arbitrary complex number we have the rule that if a state vector corresponding to a given quantum state is multiplied by any complex number (other than zero) the resulting state vector corresponds to the same quantum state. The null vector will be supposed to correspond to no state at all.

The quantum state described by a given state vector thus depends only on the "direction" or "orientation" of the state vector in the space $\mathcal{S}$. In chapter 6, when we introduced state vectors as the columns of a unitary matrix, a certain normalization condition existed for them. It is customary to impose such normalization restrictions on state vectors so as to limit the number of state vectors which one has to consider as corresponding to a given quantum state. We shall discuss these normalization restrictions later.

A state vector $|T\rangle$ is said to be dependent on, or independent of, a set of state vectors $|A\rangle$, $|B\rangle$,.... according as it can or cannot be expressed as a linear combination of the state vectors $|A\rangle$, $|B\rangle$,... as in (7.4). The quantum state corresponding to $|T\rangle$ is likewise said to be dependent on or independent of the states corresponding to $|A\rangle$, $|B\rangle$.... according as $|T\rangle$ is or is not dependent on A, B,... Evidently the members of a set of state vectors $|A\rangle$, $|B\rangle$,... which does not include the null vector (and hence the corresponding quantum states) are independent of one another if the equation $a|A\rangle + b|B\rangle + \ldots = 0$ implies $a = b = \ldots = 0$.

Denote the field of complex numbers by $\mathcal{N}$. Consider a mapping $\mathcal{S} \to \mathcal{N}$ of the state-vector space into the field of complex numbers. Such a mapping will be said to be a conjugate vector and will be denoted by the symbol $\langle\,|$ provided it satisfies the following conditions:

Denote by $\langle\,|(|A\rangle)$, or $\langle\,|A\rangle$ for short, the complex number into which the state-vector $|A\rangle$ is mapped. Then

10) $$\langle\,|(|A\rangle + |B\rangle) = \langle\,|A\rangle + \langle\,|B\rangle$$

That is, the image (under the mapping) of the sum of two state vectors is equal to the sum of their individual images.

11) $\langle | (a|A\rangle) = a \langle | A \rangle$

That is, the image of the product of a complex number with a state vector is equal to the product of the complex number with the image of the state vector.

Conjugate vectors thus defined are evidently analogous to row vectors (as distinguished from column vectors) in the theory of finite-dimensional vector spaces. The image of a given state vector $|A\rangle$ under the mapping $\langle |$ may be regarded as the scalar product of $|A\rangle$ with the conjugate vector $\langle A|$.

A particular conjugate vector may be specified by a label such as A, B, etc., thus : $\langle A|, \langle B|$, etc. The sum of two conjugate vectors, and the product of a conjugate vector with a complex number may be introduced by means of the following definitions :

12) $(\langle A| + \langle B|)|C\rangle = \langle A|C\rangle + \langle B|C\rangle$ for all C in $\mathfrak{H}$.

13) $(a\langle A|)|C\rangle = a\langle A|C\rangle$ for all C in $\mathfrak{H}$.

As a result of these definitions one may readily derive the following rules using previously stated postulates :

$$\langle A| + \langle B| = \langle B| + \langle A| \tag{7.6}$$

$$(\langle A| + \langle B|) + \langle C| = \langle A| + (\langle B| + \langle C|) \tag{7.7}$$

$$a(\langle A| + \langle B|) = a\langle A| + a\langle B| \tag{7.8}$$

$$(a+b)\langle A| = a\langle A| + b\langle A| \tag{7.9}$$

$$a(b\langle A|) = (ab)\langle A| \tag{7.10}$$

Expressions (7.7) and (7.10) may evidently, without ambiguity, be replaced by the expressions $\langle A| + \langle B| + \langle C|$ and $ab\langle A|$ respectively.

Problem XXXV : Derive equations (7.6 - 10)

The conjugate vector which maps all state-vectors into the value zero will be called the null conjugate vector and denoted by 0.

14) $\langle X|A\rangle = 0$ for all $|A\rangle$ in S implies $\langle X| = 0$

It will be seen later that no confusion will arise by denoting several different conceptual entities by the same symbol 0.

With the introduction of the convenient abbreviations

$$-\langle A| \equiv (-1)\langle A| \tag{7.11}$$

$$\langle A| - \langle B| \equiv \langle A| + (-\langle B|), \tag{7.12}$$

one may readily derive the following set of miscellaneous theorems :

$$0\langle A| = 0 \qquad \text{for all conjugate vectors } \langle A| \tag{7.13}$$

$$1\langle A| = \langle A| \qquad \text{for all conjugate vectors } \langle A| \tag{7.14}$$

$$\langle A|0 = 0 \qquad \text{for all conjugate vectors } \langle A| \tag{7.15}$$

$$a\,0 = 0 \qquad \text{for all uses of the symbol 0} \tag{7.15}$$

$$\langle A| - \langle A| = 0 \qquad \text{for all conjugate vectors } \langle A| \tag{7.17}$$

$$\langle A| + 0 = \langle A| \qquad \text{for all conjugate vectors } \langle A| \tag{7.18}$$

The symbol 0 on the left of (7.13) is the number zero while that on the right is the null conjugate vector. The symbol 0 on the left of (7.15) is the null state vector while that on the right is the number zero.

Problem XXXVI : Prove Eqs. (7.13 - 18)

Following up the analogy of state vectors with vectors in spaces of a finite number of dimensions, we shall now associate a unique conjugate

vector with every state vector. We shall write

$$\langle A| = |A\rangle^{*} \tag{7.19}$$

That is, we shall use the same label (here, A) to specify a state vector and its associated conjugate vector. The conjugate vector $\langle A|$ is frequentl called the complex conjugate of the state vector $|A\rangle$, and it is also frequently called the Hermitian adjoint, or simply the adjoint, of $|A\rangle$. The asterisk † signifies the process of taking the adjoint. The following postulates govern state-vectors and their associated adjoints.

14) $$(\,|A\rangle + |B\rangle\,)^{\dagger} = \langle A| + \langle B|$$

15) $$(a\,|A\rangle)^{\dagger} = a^{*}\langle A|$$

16) $$(\langle A|\,B\rangle\,)^{\dagger} = \langle B\,|\,A\rangle$$

The asterik on the right of 15) indicates that the ordinary complex conjugate of the number a is to be taken. Similarly, the asterik on the left of 16) indicates that the ordinary complex conjugate of the number $\langle A|\,B\rangle$ is to be taken. Without confusion we may remove the parentheses and write the expression on the left of 15) simply as $\langle A\,|\,B\rangle^{\dagger}$.

Evidently

$$\langle A|A\rangle^{*} = \langle A|A\rangle \tag{7.20}$$

That is, the number $\langle A\,|\,A\rangle$ is real. We add here the following important postulate :

17) $$\langle A\,|\,A\rangle > 0 \qquad \text{for} \qquad |A\rangle \neq 0.$$

The "lenght" of a state vector $|A\rangle$ is defined as

$$L_A = \sqrt{\langle A|A\rangle} \tag{7.21}$$

The only state vector which has zero lenght is the null vector.

The fact that the correspondence between state vectors and their complex conjugates is unique, and hence one-to-one, implies that the theory is essentially symmetrical between state-vectors and conjugate vectors, and that a quantum state can be specified by a conjugate vector just as well as by a state-vector. This means that a given state vector $|A\rangle$ may be regarded as the complex conjugate of a certain conjugate vector $\langle A|$. We shall sometimes write

$$|A\rangle = \langle A|^* \tag{7.22}$$

Also, the space of conjugate vectors will be denoted by S^*.

It is easy to show that the complex conjugate of the null state vector is the null conjugate vector. If 0 denotes the null state vector, then, by postulate 16) for any state vector $|A\rangle$

$$0^* |A\rangle = (\langle A| 0)^* = 0 \tag{7.24}$$

which implies

$$0^* = 0,$$

the 0 on the right of (7.25) denoting the null conjugate vector.

Let $|A\rangle$ and $|B\rangle$ be any two state vectors. Then, using postulates 14), 15), 16) and 17), we may write

$$\begin{aligned} 0 \leq\ & \langle B|B\rangle \left\{|A\rangle - \frac{\langle B|A\rangle}{\langle B|B\rangle}|B\rangle\right\}^* \left\{|A\rangle - \frac{\langle B|A\rangle}{\langle B|B\rangle}|B\rangle\right\} \\ =\ & \langle B|B\rangle \left\{\langle A| - \frac{\langle A|B\rangle}{\langle B|B\rangle}\langle B|\right\}\left\{|A\rangle - \frac{\langle B|A\rangle}{\langle B|B\rangle}|B\rangle\right\} \\ =\ & \langle A|A\rangle\langle B|B\rangle - \langle A|B\rangle\langle B|A\rangle \\ =\ & \langle A|A\rangle\langle B|B\rangle - |\langle A|B\rangle|^2 \end{aligned} \tag{7.26}$$

(7.26) is known as the Schwartz inequality. The angle θ between the state vectors $|A\rangle$ and $|B\rangle$ is defined by

$$\cos^2\theta = \frac{|\langle A|B\rangle|^2}{\langle A|A\rangle\langle B|B\rangle} \leq 1 \qquad (7.27)$$

The equality holds only when $|A\rangle = \frac{\langle B|A\rangle}{\langle B|B\rangle}|B\rangle$, that is, only when $|A\rangle$ and $|B\rangle$ both correspond to the same state and only differ by a numerical factor $\frac{\langle B|A\rangle}{\langle B|B\rangle}$. (If $|A\rangle = a|B\rangle$ then $\frac{\langle B|A\rangle}{\langle B|B\rangle} = a$, and this factor is seen to be arbitrary.) When the equality holds we have $\theta = 0$, and therefore we see that two state vectors are "parallel" to one another if and only if they correspond to the same quantum state. Two state vectors are "orthogonal" to one another if $\theta = \frac{\pi}{2}$ which implies

$$\langle A|B\rangle = 0. \qquad (7.28)$$

We shall often find it convenient to consider only those state vectors $|A\rangle$ corresponding to a given quantum state for which

$$\langle A|A\rangle = 1 \qquad (7.29)$$

Such state-vectors are said to be normalized. Any state vector (other than the null vector) can be normalized through multiplication by a suitable numerical factor. The process of normalization, however, does not completely determine the state-vector, as it can always be multiplied by a phase factor $e^{i\varphi}$ where φ is real. On the other hand, we shall frequently suppose that this phase factor has been chosen ahead of time and is held fixed throughout a given discussion. In this way, each quantum state will be represented by a unique state vector.

Consider now a mapping $S \to S$ of the state-vector space into itself. Such a mapping will be called a linear operator (or simply an operator) and will be denoted by a symbol such as α, β, γ, δ, ξ, etc. provided it satisfies the following conditions :

Denote by $\alpha|A\rangle$ the state vector into which $|A\rangle$ is mapped

under the mapping α . Then

18) $$\alpha(|A\rangle + |B\rangle) = \alpha|A\rangle + \alpha|B\rangle$$

19) $$\alpha(a|A\rangle = a\,\alpha|A\rangle$$

Referring to postulates 5) and 7), we see that ordinary complex numbers may be regarded as special cases of linear operators. Linear operators may be added and multiplied together through the following definitions.

20) $$(\alpha+\beta)|A\rangle = \alpha|A\rangle + \beta|A\rangle$$

21) $$(\alpha\beta)|A\rangle = \alpha(\beta|A\rangle)$$

The 1 st expression may be written, without ambiguity, simply as $\alpha\beta|A\rangle$. From postulates 18) to 21), one may readily derive the following relations

$$\alpha+\beta = \beta+\alpha \tag{7.30}$$

$$(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma) \tag{7.31}$$

$$\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma \tag{7.32}$$

$$(\alpha+\beta)\gamma = \alpha\gamma + \beta\gamma \tag{7.33}$$

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma \tag{7.34}$$

Problem XXXVII : Derive equations (7.30 -34)

Expressions (7.31) and (7.34) may, without ambiguity, be abbreviated by the expressions $\alpha+\beta+\gamma$ and $\alpha\beta\gamma$ respectively.

The operator which maps all state vectors into the null vector may be called the null operator. The null operator is, however, identical with the number zero, and there is perhaps not much point in

introducing a double terminology. The operator which maps each state-vector into itself is sometimes called the identity or the unit operator. The unit operator is identical with the number 1.

With the introduction of the abbreviations

$$-\alpha \equiv (-1)\alpha \tag{7.35}$$

$$\alpha - \beta \equiv \alpha + (-\beta) \tag{7.36}$$

one may derive the following relations

$$0\alpha = \alpha 0 = 0 \tag{7.37}$$

$$1\alpha = \alpha 1 = \alpha \tag{7.38}$$

$$\alpha - \alpha = 0 \tag{7.39}$$

$$\alpha + 0 = \alpha \tag{7.40}$$

Problem XXXVIII : Derive equations (7.37 - 40)

Starting form the above rules of combination one may construct an algebra of linear operators. In general $\alpha\beta \neq \beta\alpha$. Therefore it becomes of interest to consider bracket expressions of the form

$$[\alpha, \beta] \equiv \alpha\beta - \beta\alpha \tag{7.41}$$

$$[\alpha, \beta, \gamma] \equiv \alpha[\beta, \gamma] + \beta[\gamma, \alpha] + \gamma[\alpha, \beta] \tag{7.42}$$

$$[\alpha, \beta, \gamma, \delta] \equiv \alpha[\beta, \gamma, \delta] - \beta[\gamma, \delta, \alpha] + \gamma[\delta, \alpha, \beta] - \delta[\alpha, \beta, \gamma] \tag{7.43}$$

.... etc.

$$\{\alpha,\beta\} \equiv \alpha\beta+\beta\alpha \tag{7.44}$$

$$\{\alpha,\beta,\gamma\} \equiv \alpha\{\beta,\gamma\}+\beta\{\gamma,\alpha\}+\gamma\{\alpha,\beta\} \tag{7.45}$$

$$\{\alpha,\beta,\gamma,\delta\} \equiv \alpha\{\beta,\gamma,\delta\}+\beta\{\gamma,\delta,\alpha\}+\gamma\{\delta,\alpha,\beta\}+\delta\{\alpha,\beta,\gamma\} \tag{7.46}$$

etc....

The bracket expressions (7.41 - 43) are completely antisymmetric in each pair of operators, while the bracket expressions (7.44 - 46) are completely symmetric in each pair. By far the most important of these bracket expressions are (7.41) and (7.44). (7.41) is called the commutator bracket of α and β, and if $[\alpha,\beta] = 0$ then α and β are said to commute with one another. It is evident that complex numbers commute with all operators. (7.44) is called the anticommutator bracket of α and β, and if $\{\alpha,\beta\} = 0$, then α and β are said to anticommute witho one another.

Consider a state vector $|A\rangle$, a conjugate vector $\langle B|$ and a linear operator α. We shall denote by $\langle B|\alpha$ the conjugate vector which, for all $|A\rangle$ in $\mathcal{S}$, maps $|A\rangle$ into the same complex number as the state vector $\alpha|A\rangle$ is mapped into by the conjugate vector $\langle B|$. That is,

22) $(\langle B|\alpha)|A\rangle = \langle B|(\alpha|A\rangle)$ for all $|A\rangle$ in $\mathcal{S}$

The expression appearing in postulate 22) is conveniently written in the abbreviated form $\langle B|\alpha|A\rangle$. The quantity $\langle B|\alpha$ is called the product of the conjugate vector $\langle B|$ with the operator α. To show that $\langle B|\alpha$ is actually also a conjugate vector, one must prove that it satisfies postulates 10) and 11). That is

$$(\langle B|\alpha)(|A\rangle+|C\rangle) = (\langle B|\alpha)|A\rangle + (\langle B|\alpha)|C\rangle, \tag{7.47}$$

$$(\langle B|\alpha)(a|A\rangle) = a(\langle B|\alpha)|A\rangle, \tag{7.48}$$

for all state vectors $|A\rangle$ and $|C\rangle$ and for all complex numbers α. Equations (7.47) and (7.48) are easily verified.

Problem XXXIX : Verify equations (7.47) and (7.48)

It is now evident that a linear operator may be regarded not only as a mapping $S \to S$ of state-vector space into itself but also as a mapping $S^* \to S^*$ of conjugate-vector space into itself. Either mapping uniquely specifies the linear operator. One may readily show that products of linear operators with conjugate vectors obey rules completely analogous to postulates 18), 19), 20) and 21), namely

$$(\langle A| + \langle B|)\alpha = \langle A|\alpha + \langle B|\alpha \qquad (7.49)$$

$$(\alpha\langle A|)\alpha = \alpha\langle A|\alpha \qquad (7.50)$$

$$\langle A|(\alpha+\beta) = \langle A|\alpha + \langle A|\beta \qquad (7.51)$$

$$\langle A|(\alpha\beta) = (\langle A|\alpha)\beta \qquad (7.52)$$

Expression (7.52) may, without ambiguity, be written in the form $\langle A|\alpha\beta$. Further easily verified relations are the following :

$$\langle A|\alpha = \alpha\langle A| \quad \text{for any complex number } \alpha, \qquad (7.53)$$

$$\langle A|0 = 0 \qquad (7.54)$$

$$\langle A|1 = \langle A| \qquad (7.55)$$

Problem XL : Prove equations (7.49 - 55) :

Equation (7.53) expresses the fact that in taking the product of a conjugate vector with a complex number, the complex number may be placed either to the right or to the left of the vector. We have not so far introduced a similar convention for state-vectors, but it it will at this point be convenient to do so. Thus we write

$$|A\rangle a \equiv a|A\rangle \quad \text{for all } a, |A\rangle \tag{7.56}$$

Let $|A\rangle$ be an arbitrary state vector and $\langle B|$ an arbitrary conjugate vector. If we now introduce the notational convention

$$(|A\rangle\langle B|)|C\rangle = |A\rangle\langle B|C\rangle \tag{7.57}$$

we may regard the expression $|A\rangle\langle B|$ as a linear operator, namely, the operator which, for all state vectors $|C\rangle$, maps $|C\rangle$ into the state vector $|A\rangle\langle B|C\rangle$. It is evident that $|A\rangle\langle B|$ maps S into the set of all state-vectors parallel to $|A\rangle$. To show that $|A\rangle\langle B|$ actually is a linear operator, one must prove that it satisfies postulates 18) and 19). That is,

$$(|A\rangle\langle B|)(|C\rangle + |D\rangle) = (|A\rangle\langle B|)|C\rangle + (|A\rangle\langle B|)|D\rangle \tag{7.58}$$

$$(|A\rangle\langle B|)(a|C\rangle) = a(|A\rangle\langle B|)|C\rangle \tag{7.59}$$

for all state vectors $|C\rangle$ and $|D\rangle$ and for all complex numbers a. Equations (7.58) and (7.59) are readily verified. Also easily verified is the equation which expresses the result of multiplying $|A\rangle\langle B|$ with a conjugate vector, namely

$$\langle C|(|A\rangle\langle B|) = \langle C|A\rangle\langle B| \tag{7.60}$$

It is evident that $|A\rangle\langle B|$ maps the conjugate vector space S^* into the set of all conjugate vectors parallel to $\langle B|$.

Problem XLI : Prove equations (7.58 - 60)

Consider a state vector $|A\rangle$ and a linear operator α. We shall denote by α^* the operator which, for all $|A\rangle$ in S, maps $|A\rangle$ into the complex conjugate of the conjugate vector into which $\langle A|$ is mapped by α. That is,

$$\alpha^* |A\rangle = (\langle A|\alpha)^* \tag{7.61}$$

or, taking the complex conjugate of both sides of (7.61),

$$\langle A|\alpha = (\alpha^*|A\rangle)^* \tag{7.62}$$

From postulate 15) and relation (7.53) it is evident that if α is a complex number, α^* is its complex conjugate. More generally, if α is any linear operator, we shall call α^* its complex conjugate. α^* is also often referred to as the Hermitian adjoint of α. α^*, like α, is a linear operator, as may be verified by showing that it satisfies postulates 18) and 19). Thus

$$\begin{aligned}
\alpha^*(|A\rangle+|B\rangle) &= [(|A\rangle+|B\rangle)^*\alpha]^* && \text{by (7.61) and (7.19)}\\
&= [(\langle A|+\langle B|)\alpha]^* && \text{by postulate 14)}\\
&= (\langle A|\alpha+\langle B|\alpha)^* && \text{by (7.49)}\\
&= (\langle A|\alpha)^*+(\langle B|\alpha)^* && \text{by postulate 14) and (7.22)}\\
&= \alpha^*|A\rangle+\alpha^*|B\rangle && \text{by (7.61)} \qquad \leftarrow (7.63)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\alpha^*(a|A\rangle) &= [(a|A\rangle)^*\alpha]^* = [a^*\langle A|\alpha]^*\\
&= a(\langle A|\alpha)^* = a\alpha^*|A\rangle. \qquad (7.64)
\end{aligned}$$

Now, multiply equation (7.61) on the left by $\langle B|$ and use postulate 16). We obtain

$$\langle B|\alpha^*|A\rangle = \langle B|(\alpha^*|A\rangle) = \langle B|(\langle A|\alpha)^* = [(\langle A|\alpha)|B\rangle]^* = \langle A|\alpha|B\rangle^* \tag{7.65}$$

for all $|A\rangle$ and $|B\rangle$. From (7.65) we get

$$\langle B|\alpha^{**}|A\rangle = \langle A|\alpha^*|B\rangle^* = \langle B|\alpha|A\rangle^{**} = \langle B|\alpha|A\rangle \tag{7.66}$$

Now a state vector is completely determined by its scalar products with all conjugate vectors, just as a conjugate vector is completely specified by its scalar products with all state-vectors.

Problem XLII : Prove this statement. That is, prove that if $\langle C|A\rangle = \langle C|B\rangle$ for all $\langle C|$ then $|A\rangle = |B\rangle$. Use postulate 17).

Hence, since (7.65) is valid for all $\langle B|$ we can infer

$$\alpha^{**}|A\rangle = \alpha|A\rangle, \tag{7.67}$$

and since (7.67) is valid for all $|A\rangle$, we can finally infer

$$\alpha^{**} = \alpha \tag{7.68}$$

Let α and β be any two linear operators. Then we may write

$$\langle B|(\alpha\beta)^*|A\rangle = \langle A|\alpha\beta|B\rangle^* = (\beta|B\rangle)^*(\langle A|\alpha)^*$$
$$= (\langle B|\beta^*)(\alpha^*|A\rangle) = \langle B|\beta^*\alpha^*|A\rangle \tag{7.69}$$

Since (7.69) holds for all $|A\rangle$ and $|B\rangle$, we have

$$(\alpha\beta)^* = \beta^*\alpha^* \tag{7.70}$$

Using (7.70), we obtain

$$(\alpha\beta\gamma)^* = [\alpha(\beta\gamma)]^* = (\beta\gamma)^*\alpha^* = (\gamma^*\beta^*)\alpha^* = \gamma^*\beta^*\alpha^* \tag{7.71}$$

and, generally

$$(\alpha_1\alpha_2 \cdots \alpha_n)^* = \alpha_n^*\alpha_{n-1}^* \cdots \alpha_1^* \tag{7.72}$$

The complex conjugate of the operator $|A\rangle\langle B|$ defined by (7.57) may readily determined. For any $|C\rangle$ and $|D\rangle$, we have

$$\begin{aligned}\langle C|(|A\rangle\langle B|)^*|D\rangle &= \langle D|(|A\rangle\langle B|)|C\rangle^* = \left[\langle D|(|A\rangle\langle B|C\rangle)\right]^* \\ &= (\langle D|A\rangle\langle B|C\rangle)^* = \langle B|C\rangle^*\langle D|A\rangle^* \\ &= \langle C|B\rangle\langle A|D\rangle = \langle C|(|B\rangle\langle A|)|D\rangle\end{aligned} \tag{7.73}$$

Since (7.73) is true for all $|C\rangle$ and $|D\rangle$, we have

$$(|A\rangle\langle B|)^* = |B\rangle\langle A| \tag{7.74}$$

It is now easy to see one of most characteristic and useful features of the Dirac formalism. The complex conjugate of any product of Dirac bracket expressions is obtained by taking the complex conjugate of each factor and reversing the order of all the factors.

An operator is said to be real, or Hermitian, if it is equal to its own complex conjugate. An operator is said to be imaginary, or anti-Hermitian, if it is equal to the negative of its complex conjugate. Any operator may be split into the sum of a real and imaginary part.

Thus

$$\alpha = \beta + i\gamma \tag{7.75}$$

where

$$\beta = \frac{1}{2}(\alpha + \alpha^*) \quad \text{and} \quad \gamma = -\frac{i}{2}(\alpha - \alpha^*) \tag{7.76}$$

β and γ are both real.

If two operators, α and β, are real it is easy to see that the operators $\{\alpha, \beta\}$ and $i[\alpha, \beta]$ are also real. More generally, if the operators $\alpha_1, \alpha_2, \ldots \alpha_n$ are real, then the operators $\{\alpha_1, \ldots \alpha_n\}$ and $i^{\frac{1}{2}n(n-1)}[\alpha_1, \ldots, \alpha_n]$ are real.

Let ξ be a linear operator, and consider the equation

$$\xi |A\rangle = a |A\rangle \tag{7.77}$$

where a is an unknown number and $|A\rangle$ an unknown state-vector. When (7.77) is satisfied, with $|A\rangle \neq 0$, a is called an eigenvalue of the operator ξ and $|A\rangle$ is said to be an eigenvector of ξ corresponding to the eigenvalue a. Evidently the product of an eigenvector corresponding a certain eigenvalue with any complex number is also an eigenvector corresponding to the same eigenvalue. Later we shall be primarily interested in normalized eigenvectors.

A given linear operator may have more than one eigenvector corresponding to a given eigenvalue. In this case, any linear conbination of such eigenvectors is also an eigenvector corresponding to the same eigenvalue. It is evident that every state-vector is an eigenvector of a complex number, the eigenvalue being just the number itself.

Suppose ξ is a real operator. Then if we multiply equation (7.77) on the left by $\langle A |$ and take the complex conjugate, we get, remembering that ξ and $\langle A | A \rangle$ are real,

$$a^* \langle A|A\rangle = \langle A|\xi|A\rangle^* = \langle A|\xi|A\rangle = a\langle A|A\rangle \tag{7.78}$$

Since $\langle A | A \rangle > 0$ we may infer

$$a^* = a \tag{7.79}$$

That is, all the eigenvalues of a real linear operator are real. Taking the complex conjugate of equation (7.77) itself, we obtain

$$\langle A|\xi = a\langle A| \tag{7.80}$$

Equation (7.80) may be regarded as an eigenvector equation in the conjugate vector space S^*. Evidently, the eigenvectors of ξ in the conjugate vector space are just the complex conjugates of the eigenvectors in state-vector space, and the eigenvalues are the same in both spaces.

Dirac denotes eigenvalues of a real linear operator ξ by primed letters ξ', ξ'', ξ''', etc. The primed letters ξ', ξ'', ξ''', etc. are also used to label corresponding eigenvectors. If there is more than one eigenvector corresponding to a given eigenvalue, the various eigenvectors may be distinguished by adding extra labels, γ', γ'', γ''', etc. Thus

$$\xi|\xi',\gamma'\rangle = \xi'|\xi',\gamma'\rangle \tag{7.81}$$

Let ξ' and ξ'' be two distinct eigenvalues of a real operator ξ. If $|\xi',\gamma'\rangle$, $|\xi'',\gamma''\rangle$ are corresponding eigenvectors then we may write

$$\xi|\xi',\gamma'\rangle = \xi'|\xi',\gamma'\rangle \tag{7.82}$$

$$\xi|\xi'',\gamma''\rangle = \xi''|\xi'',\gamma''\rangle \tag{7.83}$$

Multiplying equation (7.82) on the left by $\langle\xi'',\gamma''|$ and (7.83) on the left by $\langle\xi',\gamma'|$, we obtain

$$\langle\xi'',\gamma''|\xi|\xi',\gamma'\rangle = \xi'\langle\xi'',\gamma''|\xi',\gamma'\rangle \tag{7.84}$$

$$\langle\xi',\gamma'|\xi|\xi'',\gamma''\rangle = \xi''\langle\xi',\gamma'|\xi'',\gamma''\rangle \tag{7.85}$$

Taking the complex conjugate of equation (7.84) and substracting it from

(7.85), we get

$$(\xi'' - \xi')\langle \xi', \gamma' / \xi'', \gamma'' \rangle = 0. \tag{7.86}$$

Since $\xi' \neq \xi''$, we may infer

$$\langle \xi', \gamma' / \xi'', \gamma'' \rangle = 0. \tag{7.87}$$

That is, eigenvectors corresponding to two different eigenvalues are orthogonal.

Consider an arbitrary set $\mathcal{S}$ of state-vectors. The set C will be said to be a complete set if any state-vector in $\mathcal{S}$ can be expressed as a linear combination of members of C. We shall assume that out of any complete set C one may select a complete set all of whose members are independent of one another. We might suppose the task to be accomplished somewhat as follows : First, discard the null vector if it occurs in C. Next select one remaining vector in C as choice number one. Then run through the other members of C discarding all vectors parallel to choice number one. From the remaining vectors select one as choice number two. Then run through the other remaining members of C discarding all vectors dependent on choices number one and two. From those remaining select choice number three, and so on. Continue in this manner until all the members of C are exhausted. The selected vectors will form a complete independent set. Actually, however, such a procedure cannot usually be carried out even in principle. This is because a complete set will usually contain not only an infinity of members, but a non-denumerable infinity. Nevertheless, we shall assume that complete independent sets of vectors can be obtained in this way by some sort of limiting process. A complete set of independent state vectors will be called an irreducible complete set.

From any irreducible complete set C a complete set O of mutually orthogonal non-null state vectors can be constructed. First observe that any set of mutually orthogonal state-vectors, not including the null vector, is an independent set. For, denote the members of the set by $|A\rangle, |B\rangle$.... etc. Then consider the equation

$$a|A\rangle + b|B\rangle + \ldots = 0 \tag{7.88}$$

Multiplying (7.88) on the left by $\langle A|$ we obtain

$$a\langle A|A\rangle = 0 \tag{7.89}$$

and since $\langle A|A\rangle > 0$ this implies $a = 0$. Similarly, multiplying on the left by $\langle B|$ we obtain $b = 0$, and so on. Thus $|A\rangle, |B\rangle, \ldots$ are all independent. In particular we see that any complete set of mutually orthogonal non-null state vectors is irreducible.

The set $\mathcal{O}$ may now be constructed out of the set C as follows : First arrange the members of C in some order, thus : $|A_1\rangle, |A_2\rangle, \ldots$ (If C has a non-denumerable infinity of members this ordering cannot actually be done, and one must have recourse to some sort of limiting procedure.) Then the members of $\mathcal{O}$ may be chosen to be

$$\left.\begin{aligned}
|B_1\rangle &= |A_1\rangle \\
|B_2\rangle &= |A_2\rangle - |B_1\rangle\langle B_1|A_2\rangle\langle B_1|B_1\rangle^{-1} \\
|B_3\rangle &= |A_3\rangle - |B_1\rangle\langle B_1|A_3\rangle\langle B_1|B_1\rangle^{-1} - |B_2\rangle\langle B_2|A_3\rangle\langle B_2|B_2\rangle^{-1} \\
&\cdots\cdots\cdots \\
|B_n\rangle &= |A_n\rangle - \sum_{m=1}^{n-1} |B_m\rangle\langle B_m|A_n\rangle\langle B_m|B_m\rangle^{-1}
\end{aligned}\right\} \tag{7.90}$$

It is easy to see that the $|B\rangle$'s thus constructed are all orthogonal to one another. Moreover, since the $|B\rangle$'s are simply linear combinations of the $|A\rangle$'s with at least one non-vanishing coefficient (the first), the null vector cannot be found among them. Finally, equations (7.90) can easily be solved to express the $|A\rangle$'s in terms of the $|B\rangle$'s and since C is a complete set, $\mathcal{O}$ is therefore also a complete set.

Consider the set of all eigenvectors of a real linear operator ξ. If this set is complete then ξ is said to be a <u>proper</u> real operator From the set of all eigenvectors of a proper real operator one may select an irreducible complete set C. This set will still contain eigenvectors referring to all the possible eigenvalues of ξ, since eigenvectors corresponding to different eigenvalues are, by (7.87), orthogonal and hence

independent. The members of C may be labeled by the eigenvalues ξ' together with any additional labels γ' which may be necessary, as in (7.81). From the set C one may construct an orthogonal complete set by the procedure outlined in (7.90). Let the members of the set C be divided into groups corresponding to the different eigenvalues of ξ. The method (7.90) may be applied separately to each individual group, since eigenvectors corresponding to different eigenvalues are already orthogonal. Finally, the members of the orthogonal complets set so obtained may be normalized through multiplication by suitable numerical factors; the orthogonal complete set will be called an orthonormal set. From now on we shall work with complete orthonormal sets of eigenvectors. If the labeling scheme is such that each set of labels ξ', γ' corresponds to a unique member of the orthonormal set, then we may write

$$\langle \xi', \gamma' | \xi'' \gamma'' \rangle = \delta_{\xi'\xi''} \delta_{\gamma'\gamma''} \qquad (7.91)$$

If, as is frequently the case, the labels γ' represents sets of numbers, rather than single numbers, the expression $\delta_{\gamma'\gamma''}$ must be understood as a symbolic Kronecher delta representing a product of Kronecher deltas over the various numbers involved.

The problem of finding the eigenvectors and eigenvalues of a given real operator and, in particular, of determining whether or not the operator is a proper real operator, is, in general, very difficult. It is, in fact, as we shall see, one of the central problems of quantum mechanics. There are certain cases, however, in which the problem is very easy. For example, let $|A\rangle$ be an arbitrary normalized state-vector, and consider the operator $|A\rangle\langle A|$. By (7.74) we see that this is a real operator. Moreover, it is seen that $|A\rangle$ is a eigenvector of this operator corresponding to the eigenvalue 1, while all vectors orthogonal to $|A\rangle$ are eigenvectors corresponding to the eigenvalue 0. It is evident that $|A\rangle$ may be chosen as a member of a complete orthonormal set. The members of this set are then all eigenvectors of $|A\rangle\langle A|$, the vector $|A\rangle$ corresponding to the eigenvalues 1 and all others to the eigenvalue 0. It is seen that the only eigenvalues are 0 and 1, and that $|A\rangle\langle A|$ is a proper real

operator.

Let ξ be a proper real operator. Let f be a real function which is defined over all the eigenvalues of ξ. $f(\xi)$ is defined as the linear operator which satisfies the equation

$$f(\xi)|\xi',\gamma'\rangle = f(\xi')|\xi',\gamma'\rangle \tag{7.92}$$

for all eigenvectors $|\xi',\gamma'\rangle$ of ξ. Since the eigenvectors form a complete set, $f(\xi)$ is completely defined by (7.92). For, since any state vector is expressible as a linear combination of eigenvectors the effect of multiplying any state vector by $f(\xi)$ is completely determined. The operator $f(\xi)$ is said to be a function of the operator ξ. It is evident that any eigenvector of ξ is also an eigenvector of $f(\xi)$ and that since ξ is proper, $f(\xi)$ is also proper. Moreover, it is easy to show that $f(\xi)$ like ξ, is a real operator. For, multiplying (7.92) on the left by an arbitrary conjugate eigenvector $\langle\xi'',\gamma''|$ and taking the complex conjugate we obtain

$$\begin{aligned}\langle\xi',\gamma'|f(\xi)^*|\xi'',\gamma''\rangle &= \langle\xi'',\gamma''|f(\xi)|\xi',\gamma'\rangle^* \\ &= f(\xi')\langle\xi'',\gamma''|\xi',\gamma'\rangle^* = f(\xi')\langle\xi',\gamma'|\xi'',\gamma''\rangle \\ &= \langle\xi',\gamma'|f(\xi)|\xi'',\gamma''\rangle\end{aligned}$$

(7.93)

Since (7.93) holds for any pair of eigenvectors $|\xi',\gamma'\rangle, |\xi'',\gamma''\rangle$ we may infer

$$f(\xi)^* = f(\xi) \tag{7.94}$$

Let ξ' be any eigenvalue of ξ and consider tje function $\delta_{\xi'\xi}$. We have

$$\delta_{\xi'\xi}|\xi'',\gamma''\rangle = \delta_{\xi'\xi''}|\xi'',\gamma''\rangle. \tag{7.95}$$

But

$$\sum_{\xi',\alpha'} |\xi',\alpha'\rangle\langle\xi',\alpha'|\xi'',\alpha''\rangle = \sum_{\xi',\alpha'} \delta_{\xi'\xi''}\delta_{\alpha'\alpha''}|\xi',\alpha'\rangle$$
$$= \delta_{\xi'\xi''}|\xi'',\alpha''\rangle = \delta_{\xi'\xi''}|\xi'',\alpha''\rangle \qquad (7.96)$$

Hence

$$\delta_{\xi'\xi''} \equiv \sum_{\alpha'} |\xi',\alpha'\rangle\langle\xi',\alpha'| \qquad (7.97)$$

Now, also

$$\sum_{\xi'} \delta_{\xi'\xi}|\xi'',\alpha''\rangle = \sum_{\xi'}\delta_{\xi'\xi''}|\xi'',\alpha''\rangle = |\xi'',\alpha''\rangle \qquad (7.98)$$

which implies

$$\sum_{\xi'} \delta_{\xi'\xi} = 1 \qquad (7.99)$$

or

$$\sum_{\xi',\alpha'} |\xi',\alpha'\rangle\langle\xi',\alpha'| = 1 \qquad (7.100)$$

Equations (7.97) and (7.100) depend, of course, on the fact that ξ is a proper real operator and hence that the eigenvectors $|\xi',\alpha'\rangle$ form a complete set. By means of (7.100) one may immediatly obtain the expansion of an arbitrary state-vector $|A\rangle$ in terms of the eigenvectors of ξ. We have

$$|A\rangle = 1|A\rangle = \sum_{\xi',\alpha'} |\xi',\alpha'\rangle\langle\xi',\alpha'|A\rangle \qquad (7.101)$$

If the eigenvectors of ξ did not form a complete set such an expansion would not be possible. The expansion is, of course, unique. An arbitrary conjugate vector may also be expanded in the same way. Thus

$$\langle A| = \sum \langle A|\xi',\alpha'\rangle\langle\xi',\alpha'|. \qquad (7.102)$$

Let ξ and η be any two proper real operators which commute with each other. Let the eigenvectors of ξ and η be denoted respectively by $|\xi',\delta'\rangle$ and $|\eta',\delta'\rangle$. We may expand any $|\eta',\delta'\rangle$ in the form

$$|\eta',\delta'\rangle = \sum_{\xi',\delta'} |\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle \tag{7.103}$$

Applying the operator $\eta - \eta'$ to the left of equation (7.103), we obtain

$$0 = \sum_{\xi',\delta'} (\eta-\eta')|\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle. \tag{7.104}$$

But

$$\xi(\eta-\eta')|\xi',\delta'\rangle = (\eta-\eta')\xi|\xi',\delta'\rangle$$

so that $(\eta-\eta')|\xi',\delta'\rangle$ is seen to be an eigenvecotr of ξ corresponding to the eigenvalue ξ'. This means that

$$\langle\xi'',\delta''|(\eta-\eta')|\xi',\delta'\rangle = 0 \quad \text{for } \xi' \neq \xi''. \tag{7.105}$$

Hence, multiplying the equation (7.104) on the left by the operator of (7.100), we obtain

$$0 = \sum_{\xi',\delta';\xi'',\delta''} |\xi'',\delta''\rangle\langle\xi'',\delta''|(\eta-\eta')|\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle$$

$$= \sum_{\xi',\delta';\delta''} |\xi',\delta''\rangle\langle\xi',\delta''|(\eta-\eta')|\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle \tag{7.106}$$

which implies

$$\sum_{\delta'} \langle\xi',\delta''|(\eta-\eta')|\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle = 0. \tag{7.107}$$

Since equation (7.107) holds for all ξ',δ'' we may infer

$$\sum_{\delta'} (\eta-\eta')|\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle = 0. \tag{7.108}$$

That is, $\sum_{\delta'} |\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle$ is an eigenvector of η corresponding to the eigenvalue η' as well as being a simultaneous eigenvector of ξ corresponding to the eigenvalue ξ'. We may write

$$\sum_{\delta'} |\xi',\delta'\rangle\langle\xi',\delta'|\eta',\delta'\rangle = c_{\xi',\eta',\delta'}|\xi',\eta',\delta'\rangle \qquad (7.109)$$

where the coefficient $c_{\xi',\eta',\delta'}$ is to be chosen so that the vector $|\xi',\eta',\delta'\rangle$ is normalized. Evidently we must have

$$|c_{\xi',\eta',\delta'}|^2 = \sum_{\delta'',\delta'''} \langle\eta',\delta'|\xi',\delta''\rangle\langle\xi',\delta''|\xi',\delta'''\rangle\langle\xi',\delta'''|\eta',\delta'\rangle$$

$$= \sum_{\delta'} |\langle\xi',\delta'|\eta',\delta'\rangle|^2. \qquad (7.110)$$

We may now write equation (7.103) in the form

$$|\eta',\delta'\rangle = \sum_{\xi'} c_{\xi',\eta',\delta'}|\xi',\eta',\delta'\rangle. \qquad (7.111)$$

Since any state-vector can be expanded in terms of the eigenvectors $|\eta',\delta'\rangle$, it can evidently also be expanded in terms of the simultaneous eigenvectors $|\xi',\eta',\delta'\rangle$ of ξ and η. Thus the simultaneous eigenvectors of any two commuting proper real operators form a complete set.

It should be noted that the set of vectors $|\xi',\eta',\delta'\rangle$ is not necessarily irreducible. For example, the vectors $|\eta',\delta'\rangle$ might have been chosen initially to be already simultaneous eigenvectors of ξ, in which case the labels δ' would have included ξ', and some of the $c_{\xi',\eta',\delta'}$ (those for which δ' did not correspond to ξ') would have vanished. This redundancy in the labels δ' is a characteristic feature and is to be expected. After having obtained the simultaneous eigenvectors $|\xi',\eta',\delta'\rangle$ we shall still, in general, have the task of extracting from them an irreducible orthonormal set.

The converse of the above result is also true, namely, if ξ and η are any two proper real operators whose simultaneous eigenvectors form a complete set then ξ and η commute. For, if the simultaneous

[illegible] [illegible]

eigenvectors are denoted by $|\xi', \eta', \delta'\rangle$ (the labels δ' now being non-redundant), then

$$[\xi, \eta]\,|\xi', \eta', \delta'\rangle = (\xi\eta - \eta\xi)\,|\xi', \eta', \delta'\rangle = 0 \quad (7.112)$$

Since (7.112) holds for all $|\xi', \eta', \delta'\rangle$ we have

$$[\xi, \eta] = 0. \quad (7.113)$$

The above results may be readily generalized to the cases of three or more mutually commuting proper real operators. Let $\xi_1 \ldots \xi_n$ be a set of n mutually commuting proper real operators. Their simultaneous eigenvectors may be denoted by $|\xi_1' \ldots \xi_n', \delta'\rangle$, where the δ' are any additional lables which may be necessary in order to distinguish all the members of an orthonormal set. Just as we were able to define an arbitrary function of a single proper real operator, so may we define functions of several mutually commuting proper real variables. Let f be a real function of n real variables which is defined over all the eigenvalues of the n operators $\xi_1 \ldots \xi_n$. Then $f(\xi_1 \ldots \xi_n)$ is defined as the linear operator which satisfies the equation

$$f(\xi_1 \ldots \xi_n)\,|\xi_1' \ldots \xi_n', \delta'\rangle = f(\xi_1' \ldots \xi_n')\,|\xi_1' \ldots \xi_n', \delta'\rangle \quad (7.114)$$

for all $|\xi_1' \ldots \xi_n', \delta'\rangle$. Just as in the case of a function of a single proper real operator, we may easily show that $f(\xi_1 \ldots \xi_n)$ is both real and proper.

Consider a complete orthonormal set θ of state-vectors $|\lambda_1 \ldots \lambda_n\rangle$, each labeled by an ordered set of n real numbers $\lambda_1 \ldots \lambda_n$ in such a way that for each member of θ there corresponds a unique label. Then since θ is orthonormal we may write

$$\langle \lambda_1' \ldots \lambda_n' | \lambda_1'' \ldots \lambda_n'' \rangle = \delta_{\lambda_1'\lambda_1''} \cdots \delta_{\lambda_n'\lambda_n''}. \quad (7.115)$$

For m = 1,2 .... n, define L_m to be the linear operator for which

$$L_m |\lambda_1 \ldots \lambda_n\rangle = \lambda_m |\lambda_1 \ldots \lambda_n\rangle \qquad (7.116)$$

for all $|\lambda_1 \ldots \lambda_n\rangle$. Multiplying (7.116) on the left by $\langle\lambda'_1 \ldots \lambda'_n|$ and taking the complex conjugate, we obtain

$$\langle\lambda_1 \ldots \lambda_n | L_m^* | \lambda'_1 \ldots \lambda'_n\rangle = \langle\lambda'_1 \ldots \lambda'_n | L_m | \lambda_1 \ldots \lambda_n\rangle^*$$
$$= \lambda_m \langle\lambda'_1 \ldots \lambda'_n | \lambda_1 \ldots \lambda_n\rangle^* = \lambda_m \langle\lambda_1 \ldots \lambda_n | \lambda'_1 \ldots \lambda'_n\rangle$$
$$= \langle\lambda_1 \ldots \lambda_n | L_m | \lambda'_1 \ldots \lambda'_n\rangle \qquad (7.117)$$

Since (7.117) holds for all $\langle\lambda_1 \ldots \lambda_n|$ and $|\lambda'_1 \ldots \lambda'_n\rangle$ we may infer that $L_m^* = L_m$. That is, the operators L_m are real. The L_m are also seen to be proper and to possess a complete set of simultaneous eigenvectors, namely the $|\lambda_1 \ldots \lambda_n\rangle$. Therefore the L_m all commute with one another :

$$[L_m, L_r] = 0 \quad \text{for all } m, r. \qquad (7.118)$$

A set of mutually commuting proper real operators will be said to be complete if each of their simultaneous eigenvectors corresponds to a unique set of eigenvalues. The operators L_m above evidently form a complete set. If we have any set of mutually commuting operators $\xi_1 \ldots \xi_n$ which is not complete, we can add operators $L_1, L_2, \ldots$ to this set so that it becomes complete. For example, the labels γ' appearing in equation (7.114) may be treated like the labels $\lambda_1 \ldots \lambda_n$ above, and extra operators corresponding to this labels may be introduced by equations of the form (7.116). From now on we shall suppose (unless stated to the contrary) that the sets of communiting operators, $\xi_1, \ldots \xi_n$, with which we deal are complete sets. The eigenvectors will then be denoted simply by $|\xi'_1 \ldots \xi'_n\rangle$ and the orthonormality and completeness conditions will be

expressed respectively by

$$\langle \xi'_1 \ldots \xi'_n | \xi''_1 \ldots \xi''_n \rangle = \delta_{\xi'_1 \xi''_1} \ldots \delta_{\xi'_n \xi''_n} \qquad (7.119)$$

and

$$\sum_{\xi'_1 \ldots \xi'_n} | \xi'_1 \ldots \xi'_n \rangle \langle \xi'_1 \ldots \xi'_n | = 1 . \qquad (7.120)$$

For many purposes it will be convenient to introduce an abbreviation whereby we use a single symbol ξ' to stand for the ordered set of real numbers $\xi'_1 \ldots \xi'_n$. We shall then write eqs. (7.119 - 120) in the symbolic forms

$$\langle \xi' | \xi'' \rangle = \delta_{\xi' \xi''} \qquad (7.121)$$

$$\sum_{\xi'} | \xi' \rangle \langle \xi' | = 1 \qquad (7.122)$$

It must always be remembered in these and similar expressions, however, that ξ' stands not for a single number, but rather for an ordered set of numbers.

A few remarks should be made concerning difficulties which will arise when the above formalism is applied to actual practical examples. First of all, it may happen that the number, n, of individual labels which is required in order to specify a particular member of a complete orthonormal set of state-vectors, is infinite. The number may even be non-denumerably infinite. This will be the typical situation, for example, in the quantum theory of fields. It will nevertheless usually turn out, that this situation will cause no difficulty since, for the eigenvectors of interest, all but a finite number of the labels have fixed constant values, usually zero. If such is not the case, however, then, with the use of the above formalism, the description of the physical situation can only be approached by means of some sort of limiting process.

Secondly, it will very often happen that summations, such as occur for example in eqs. (7.96 - 98), (7.100 -104), (7.106 - 111),

(7.120), (7.122) cannot actually be carried out because the values which one or more of the labels can assume cover a continuous range. It is possible to circumvent this difficulty by introducing integrals as well as sums. This is very commonly done in the basic mathematical theory itself (e.g. Dirac). We shall, however, leave the summations as they stand. This procedure has the advantage of leaving the basic theory in as uniform and simple a form as possible, and it will not really lead to difficulty. For, in cases in which the labels vary over continuous ranges the ordinary summation formalism can still be applied with the aid of suitable limiting processes. In this way the basic theory will lead us by a uniform path in all cases to the mathematical formulation of the various physical problems under consideration. The integrals will appear only in the mathematical analysis of the individual physical systems. That is, although the basic theory lacks complete rigor in connection with illegitimate summations, it will be seen as we go along that it will always lead unambiguously to completely rigorous formulations of specific physical problems.

We are now ready actually to discuss the physical significance of the various concepts which we have introduced in this chapter. Consider a complete set of commuting proper real operators ξ, together with the complete orthonormal set of their simultaneous eigenvectors $|\xi'\rangle$. Let $|\ \rangle$ be an arbitrary state vector, and consider the set of all numbers $\langle\xi'|\ \rangle$. If these numbers be arranged into a one-column array $(\langle\xi'|\ \rangle)$ then we have an object like the column vectors considered in the preceding chapter. Let α be a linear operator, and consider the analogous array formed from the vector $\alpha|\ \rangle$, namely $(\langle\xi'|\alpha|\ \rangle)$. Using the completeness condition (7.122), we may express the individual elements of this latter array in the form

$$\langle\xi'|\alpha|\ \rangle = \langle\xi'|(\alpha\cdot 1)|\ \rangle = \sum_{\xi''}\langle\xi'|\alpha|\xi''\rangle\langle\xi''|\ \rangle \qquad (7.123)$$

That is, the latter array may be obtained from the previous array through multiplication by the matrix $(\langle\xi'|\alpha|\xi''\rangle)$. An operator is thus seen to be representable by a matrix, the choice of a particular complete ortho-

normal set of vectors $|\xi'\rangle$ being said to correspond to a particular representation. The vectors $|\xi'\rangle$ are called the basic vectors of the representation, the whole set of them being called a basis.

The matrix corresponding to the product of two operators α and β, in the representation of the $|\xi'\rangle$'s is given by

$$\langle \xi' | \alpha\beta | \xi'' \rangle = \langle \xi' | \alpha \times 1 \times \beta | \xi'' \rangle = \sum_{\xi'''} \langle \xi' | \alpha | \xi''' \rangle \langle \xi''' | \beta | \xi'' \rangle \tag{7.124}$$

Thus the matrix of the product of two operators is seen to be the product of the matrices of the two operators.

The process of passing from one representation to another is very easy to describe in terms of the formalism of the present chapter. Consider a complete set of operators η distinct from the operators ξ. In terms of their corresponding normalized eigenvectors $|\eta'\rangle$, we may write

$$\langle \eta' | \alpha | \eta'' \rangle = \langle \eta' | (1 \times \alpha \times 1) | \eta'' \rangle = \sum_{\xi', \xi''} \langle \eta' | \xi' \rangle \langle \xi' | \alpha | \xi'' \rangle \langle \xi'' | \eta'' \rangle \tag{7.125}$$

But

$$\sum_{\xi'} \langle \eta' | \xi' \rangle^* \langle \eta'' | \xi' \rangle = \sum_{\xi'} \langle \eta'' | \xi' \rangle \langle \xi' | \eta' \rangle$$

$$= \langle \eta'' | \eta' \rangle = \delta_{\eta'\eta''} \tag{7.126}$$

and

$$\sum_{\eta'} \langle \eta' | \xi' \rangle^* \langle \eta' | \xi'' \rangle = \sum_{\eta'} \langle \xi' | \eta' \rangle \langle \eta' | \xi'' \rangle$$

$$= \langle \xi' | \xi'' \rangle = \delta_{\xi'\xi''}. \tag{7.127}$$

Hence the matrix $(\langle \eta' | \xi' \rangle)$ is a unitary matrix, $(\langle \xi'' | \eta' \rangle)$ is its inverse, and the transformation (7.125) is a unitary transformation. That is, the passage from one basis (or representation) to another is effected by means of a unitary transformation.

The matrix relations which we have presented here are clearly just like those introduced in chapter 6. There we suggested that the quan-

tum analog of every classical observable should be regarded as an operator, its expression in any representation taking the form of a matrix. Accordingly, we shall now postulate that to every classical observable there corresponds a linear operator in the quantum theory. The question immediately arises as to what sort of operators we must select as quantum operators. Let us first observe that all individual classical observables are essentially real quantities. One might suppose that a classical observable could be a complex quantity, but it should be remembered that the observation of a complex quantity would really entail individual measurements of its real and imaginary parts. Thus we lose no generality in considering only real classical quantities. In chapter 6, equation (6.36), we have seen that the quantum matrix corresponding to a real classical quantity F must be an Hermitian matrix. This means that the corresponding quantum operator must be real. For, we have

$$\langle \xi' | F | \xi'' \rangle = \langle \xi'' | F | \xi' \rangle^* = \langle \xi' | F^* | \xi'' \rangle \qquad (7.128)$$

Since eq. (7.128) holds for all $\langle \xi' |, | \xi'' \rangle$, we have

$$F^* = F \qquad (7.129)$$

Perhaps the most important classical observable is the energy or total Hamiltonian function of a system. All the matrices considered in chapter 6. were defined with respect to bases composed of eigenvectors of unspecified (perturbed or unperturbed) energy operators. It was pointed out there, however, that there was no fundamental reason for singling out energy operators as quantities with respect to which to form bases, and that we should ultimately want to generalize the theory so that all operators are put on a more or less equal footing. Therefore we shall postulate that our bases can be defined with respect to any set of observables for which the corresponding operators form a complete commuting set. This means that in addition to requiring quantum operators to be real, we must also require them to be proper. Actually, the determination of whether or not a given

operator is proper is, in general, a very difficult, if not impossible, task. Whether or not an operator is proper depends ultimately on what kind of algebraic or analytic relations the operator satisfies, and these relations will, in turn, be seen to depend on the dynamical properties of the system in which the operators are defined. These dynamical properties are determined by the Hamiltonian function of the system. Hence, the key to whether or not our assumption, that the quantum operators are proper, is valid, really lies hidden in the Hamiltonian functions of various physical systems. Only the operators belonging to certain simple systems or certain simple though general types of operators (such as those corresponding to generalized coordinates and their conjugate momenta) have been proved to be proper. Our procedure in practice therefore will be to assume in advance that all the observables of interest correspond to proper operators, and then to proceed from there. If we ever run into a contradiction on this basis, for some particular system, we shall have to stop right there and make a very special study of the system in question. For, a fundamental modification of our whole theory will then be necessary. Fortunately this situation has not yet arisen in physics.

In chapter 6 the energy eigenvalues corresponding to various energy state-vectors were assumed to be all distinct. That is, the physical systems considered were assumed to be all non-degenerate. As was pointed out there, this really meant no loss of generality as far as the theory of states is concerned, since a degenerate system can be treated as a limiting case in which certain small degeneracy-removing perturbations tend to zero. However, when practical problems have to be worked out this procedure can be seriously defective, and, in any case, it is a rather artificial limitation to put on the formalism. The case of degeneracy can be handled very simply with the formalism of the present chapter. If the energy eigenvalues do not specify all the eigenvectors uniquely, it is only necessary to add further real operators which commute with the energy operator, and with each other, until one obtains a complete set; each simultaneous eigenvector of these operators will then be uniquely determined by the various eigenvalues.

The statistical nature of quantum mechanics has been repeatedly pointed out in preceding chapters. We must now make some postulational statement which will enable us to interpet the various quantities appearing in the foregoingformalism in statistical terms, and hence ultimately in terms of physical measurements. We shall do this simply by generalizing equation (6.63) of the preceding chapter. We may, without causing confusion, denote an observable and its corresponding quantum operator by the same symbol. We shall also denote a quantum state and its corresponding state-vector by the same symbol. Let a system be in a certain state $|\ \rangle$. Then the average or expectation value of an observable α in that state will be given by

$$\bar{\alpha} = \langle\ |\alpha|\ \rangle \qquad (7.130)$$

When the observable α is actually measured, of course, the result of the measurement can only be one of the permissable quantum values of α. As was suggested in chapter 6, these permissable values will be taken to be the eigenvalues of the operator α. It now becomes of interest to ask what the relative probabilities are that one or another of the permissable values will actually turn out to be the measured value.

Let α' be one of the eigenvalues of the operator α. Consider the following function of α: $\delta_{\alpha'\alpha}$. In the classical theory this function is automatically measured to be zero whenever α is measured to have a value different from α', and unity whenever α is measured to have the value α'. By the Correspondance Principle the same must be true in the quantum theory. If one carries out a long series of experiments repeatedly measuring the same observable α with the system initially always in the same state, then one can obtain the probability that the measured value will, in future identical experiments, be α', by chalking up a 1 whenever α occurs and a zero when it does not occur, adding up all the 1's and dividing by the total number of measurements. This, of course, is just the average or expectation value of the observable $\delta_{\alpha\alpha'}$. If we denote the normalized eigenvectors of α by $|\alpha',\gamma'\rangle$, where the γ

are any additional operators necessary to make a complete set, then, with the aid of equation (7.97) we may write this expectation value in the form

$$\langle |\delta_{\alpha'\alpha}| \rangle = \sum_{\gamma'} \langle |\alpha',\gamma'\rangle\langle\alpha',\gamma'| \rangle$$

$$= \sum_{\gamma'} |\langle \alpha',\gamma'| \rangle|^2 \qquad (7.131)$$

If α forms a complete set by itself, or, if the measurement consists of the simultaneous measurement of a complete set of observables, so that "α" is a symbol denoting a complete set of operators, equation (7.131) becomes simply

$$\langle |\delta_{\alpha'\alpha}| \rangle = |\langle \alpha'| \rangle|^2. \qquad (7.132)$$

That is, the probability that the result of the measurement will be α' is just the square of the absolute value of the element corresponding to α' in the one-column array representing the vector $|\ \rangle$ in the representation formed by the basic vectors $|\alpha'\rangle$. This element, namely $\langle\alpha'|\rangle$ is also the coefficient of $|\alpha'\rangle$ in the expansion of $|\ \rangle$ in terms of the $|\alpha'\rangle$:

$$|\ \rangle = \sum_{\alpha'} |\alpha'\rangle\langle\alpha'|\ \rangle. \qquad (7.133)$$

The coefficients $\langle\alpha'|\ \rangle$ are frequently referred to as probability amplitudes. We may readily write the expectation value of α itself in terms of these probability amplitudes.

$$\bar{\alpha} = \langle |\alpha| \rangle = \sum_{\alpha',\alpha''} \langle |\alpha'\rangle\langle\alpha'|\alpha|\alpha''\rangle\langle\alpha''|\ \rangle$$

$$= \sum_{\alpha'\alpha''} \langle |\alpha'\rangle\alpha''\langle\alpha'|\alpha''\rangle\langle\alpha''|\ \rangle$$

$$= \sum_{\alpha'} \langle |\alpha'\rangle\alpha'\langle\alpha'|\ \rangle = \sum_{\alpha'} \alpha' |\langle\alpha'|\ \rangle|^2 \qquad (7.134)$$

This is the expression we should expect, of course. The squares of the absolute values of the probability amplitudes are the weight factors in the sum over α'.

Of frequent interest is also the mean square deviation of α from its expectation value. This is defined in quantum theory, just as in classical statistical theory, by

$$\Delta\alpha^2 = \overline{(\alpha-\bar{\alpha})^2} = \overline{(\alpha^2 - 2\alpha\bar{\alpha} + \bar{\alpha}^2)}$$
$$= \overline{\alpha^2} - 2\bar{\alpha}\bar{\alpha} + \bar{\alpha}^2 = \overline{\alpha^2} - \bar{\alpha}^2 \qquad (7.135)$$

The last form is obtained by using the fact that the expectation value of a sum is clearly the sum of the expectation values, and the expectation value of a product with a number is the product of the expectation value with the number.

Before rewriting equation (7.135) in terms of probability amplitudes we should say a few words about diagonal matrices. It is evident that the matrix corresponding to any observable in a representation which is defined with respect to that observable (plus any other observables which are necessary to make up a complete set) is diagonal. Thus

$$\langle\alpha'|\alpha|\alpha''\rangle = \alpha''\langle\alpha'|\alpha''\rangle = \alpha'\langle\alpha'|\alpha''\rangle$$
$$= \alpha''\delta_{\alpha'\alpha''} = \alpha'\delta_{\alpha'\alpha''}. \qquad (7.136)$$

Moreover, from the definition (7.92), it is evident that the matrix corresponding to any function of α is also diagonal in this representation. If α denotes a complete set of observables then the fact that these observables must all be mutually commuting is reflected in the fact that the matrices corresponding to these observables in the above representation are all diagonal, and diagonal matrices always commute with one another. The process of choosing a basis defined with respect to a complete set of commuting operators is often referred to as diagonalizing the matri-

ces associated with these operators, or, loosely, as diagonalizing the operators themselves.

Any operator is defined by its matrix in a given representation. In particular, the operator α^2 is defined by the matrix

$$\langle \alpha' | \alpha^2 | \alpha'' \rangle = \alpha'^2 \delta_{\alpha'\alpha''} = \alpha''^2 \delta_{\alpha'\alpha''} \tag{7.137}$$

We may evidently write

$$\overline{\alpha^2} = \langle | \alpha^2 | \rangle = \sum_{\alpha', \alpha''} \langle | \alpha' \rangle \langle \alpha' | \alpha^2 | \alpha'' \rangle \langle \alpha'' | \rangle$$

$$= \sum_{\alpha'} \alpha'^2 \left| \langle \alpha' | \rangle \right|^2 \tag{7.138}$$

which is just the form we should expect. More generally,

$$\overline{f(\alpha)} = \sum_{\alpha'} f(\alpha') \left| \langle \alpha' | \rangle \right|^2 \tag{7.139}$$

If α denotes an ordered set of operators $\alpha_1, \ldots, \alpha_n$ the symbol α^2 may be regarded as standing for the ordered set of operators $\alpha_1^2, \ldots, \alpha_n^2$ and the symbol α'^2 as standing for the ordered set of numbers $\alpha_1'^2, \ldots, \alpha_n'^2$.

If the state-vector $| \rangle$ happens to be an eigenvector $| \alpha' \rangle$ of α then

$$\overline{\alpha} = \sum_{\alpha''} \alpha'' \left| \langle \alpha'' | \alpha' \rangle \right|^2 = \alpha' \tag{7.140}$$

and

$$\overline{\alpha^2} = \sum_{\alpha''} \alpha''^2 \left| \langle \alpha'' | \alpha' \rangle \right|^2 = \alpha'^2 \tag{7.141}$$

so that

$$\Delta\alpha^2 = \overline{\alpha^2} - \overline{\alpha}^2 = \alpha'^2 - \alpha'^2 = 0, \tag{7.142}$$

This vanishing of the mean square deviation means that the result of the measurement is certain to be α'.

That is, whenever a system is in a state which corresponds to an eigenvector of a given operator, a measurement of the observable corresponding to that operator is certain to give as a result the eigenvalue in question. The state of the system is then said to be an eigenstate of the operator or of the corresponding observable.

Referring back to postulate 16) we see that for any two representations, in terms of ξ's and η's say, we may write

$$|\langle \xi' | \eta' \rangle|^2 = |\langle \eta' | \xi' \rangle|^2 \qquad (7.143)$$

As a result of the statistical interpretations of our formalism, which have been given above, this simple equation may be read off as a fundamental theorem, known as the theorem of reprocity :

> The probability of the ξ's having the values ξ' in the state for which the η's certainly have the values η' is equal to the probability of the η's having the values η' in the state for which the ξ's certainly have the values ξ'.

Consider two observables, α and β, of some physical system. Suppose that the system is in the quantum state $|\rangle$. Then we have seen that the expectation values of these observables are given by

$$\overline{\alpha} = \langle |\alpha| \rangle , \quad \overline{\beta} = \langle |\beta| \rangle . \qquad (7.144)$$

We shall now obtain an important relation connecting the mean square deviations of the measured values from these average values in the state $|\rangle$. From (7.135) we have

$$\Delta\alpha^2 = \langle |(\alpha - \overline{\alpha})^2| \rangle , \quad \Delta\beta^2 = \langle |(\beta - \overline{\beta})^2| \rangle \qquad (7.145)$$

Since α and β are real we may write these expressions in the form

$$\Delta\alpha^2 = \langle A|A\rangle \quad , \quad \Delta\beta^2 = \langle B|B\rangle \tag{7.146}$$

where

$$|A\rangle = (\alpha - \bar{\alpha})|\rangle, \quad |B\rangle = (\beta - \bar{\beta})|\rangle. \tag{7.147}$$

Now, using the Schwartz inequality (7.26), we may write

$$\Delta\alpha^2 \Delta\beta^2 = \langle A|A\rangle \langle B|B\rangle \geqslant |\langle A|B\rangle|^2, \tag{7.148}$$

or, in terms of the root mean square deviations,

$$\begin{aligned}
\Delta\alpha\Delta\beta &\geqq |\langle A|B\rangle| = |\langle |(\alpha-\bar{\alpha})(\beta-\bar{\beta})|\rangle| \\
&= |\langle |(\tfrac{1}{2}\{\alpha,\beta\} + \tfrac{1}{2}[\alpha,\beta] - \bar{\alpha}\bar{\beta})|\rangle| \\
&= |\langle |(\tfrac{1}{2}\{\alpha,\beta\} - \bar{\alpha}\bar{\beta})|\rangle + \tfrac{i}{2}\langle |-i[\alpha,\beta]\rangle| \\
&= \left[\langle |(\tfrac{1}{2}\{\alpha,\beta\} - \bar{\alpha}\bar{\beta})|\rangle^2 + \tfrac{1}{4}\langle |-i[\alpha,\beta]|\rangle^2\right]^{1/2} \\
&\geqq \tfrac{1}{2}\langle |-i[\alpha,\beta]|\rangle .
\end{aligned} \tag{7.149}$$

In passing to the next to the last line above we made use of the fact that the operators $\frac{1}{2}\{\alpha,\beta\} - \bar{\alpha}\bar{\beta}$ and $-i[\alpha,\beta]$ are real, and hence that their expectation values are real. Equation (7.149) shows that if $[\alpha,\beta] \neq 0$ then the root mean square deviations of the operators cannot in general, vanish simultaneously. This fact has several important consquences. First of all, it indicates what we have already seen, namely, that two proper real operators which do not commute with each other cannot possess

a complete set of simultaneous eigenvectors. Secondly, suppose one makes a measurement of the observable α and obtains a precisely defined result α'. Then, in order that physics be a consistent science and not completely chaotic, one must assume that the system is subsequently in an eigenstate $|\alpha', \gamma'\rangle$ of α with $\Delta\alpha = 0$, at least for a period of time short enough so that the dynamical variation of the system does not cause α to vary appreciably. This means that if we immediately follow up our measurement of α with another measurement of α, we should get the same result α'. If, on the other hand, we immediately follow up our measurement of α with a measurement of β, obtaining a precise result β', then we know that we have caused the state of the system to become an eigenstate $|\beta'\delta'\rangle$ of β, with $\Delta\beta = 0$. The inequality (7.149) then tells us that we can, in general, no longer have $\Delta\alpha = 0$, and the state can no longer be an eigenstate of α. In fact we shall have $\Delta\alpha = \infty$, so that the value of α will be completely indeterminate. Thus the two measurements interfere with one another. This interference is a characteristic feature of the quantum theory and represents a fundamental limitation on the delicacy with which we can perform experiments. This limitation does not exist in classical theory where we can, in principle, make our measurements as delicate as we wish, so as to avoid interference effects.

Having pointed out the existence of interference effects in measurements, we should now go back and take another look at our general theory. We have spoken of the observables corresponding to a complete set of commuting proper real operators ξ as being "simultaneously measurable" (see p.35, following eq. (7.131). Perhaps now we should be a little more cautious, supposing that interference effects might prevent one from being able to set up an experiment which really makes a simultaneous measurement of all, or even several, of the observables. (One thing at a time, shall we say?) Very well then, we may still make individual measurements on the individual observables in rapid succession one after another. The relation (7.149) now imposes no restriction on the smallness of the root mean square deviations $\Delta\xi$. As a result of the successive measurements the $\Delta\xi'$s may all be set equal to zero one after another. Now, a simultaneous measurement may clearly be regarded as the limit of a set of succes-

COURS DE L'ÉCOLE D'ÉTÉ DE PHYSIQUE THÉORIQUE

QUANTUM MECHANICS

Bryce S. DE WITT

Radiation Laboratory
University of California

★ ★ ★ ★

LES HOUCHES
ÉTÉ 1953

Q U A N T U M M E C H A N I C S

Lecture Notes by

Bryce S. DE WITT
Radiation Laboratory
University of California

Chapters VIII, IX, X.

Université de Grenoble
Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

8.- GENERAL QUANTUM DYNAMICAL THEORY.

As we have already remarked in chapter 6, quantum dynamical theory is to be based on analogy with classical dynamical theory. A most important concept used in connection with classical dynamical theory is that of canonical transformations. We have seen in chapter 6 how the Correspondence Principle leads us to identify canonical transformations in classical theory with unitary transformations in quantum theory. In particular, we are led to associate the Poisson brackets which appear in expressions for infinitesimal canonical transformations with commutator brackets in the quantum theory. We shall now take the trouble to rederive these results in a slightly more rigorous and logical fashion.

Consider a set of proper real operators A, B, C,... corresponding to certain classical observables. Let these operators stand in a product in the form

$$F = ABC \ldots \tag{8.1}$$

Now suppose that the operators A, B, C, ... suffer infinitesimal variations δA, δB, δC,..... The corresponding variation in the product is given by

$$\delta F = \delta A\, BC \ldots + A\, \delta B\, C \ldots + AB\, \delta C \ldots + \ldots \tag{8.2}$$

Unlike the situation in classical theory, the various factors in the terms above must be left in the order in which they stand, since operators, in general, do not commute with one another.

Suppose the variations δA , δB , δC , ... are those produced by an infinitesimal canonical transformation in the corresponding classical observables. Then in the classical theory we may write (referring to eq. (1.56)),

$$\delta A = \varepsilon (A,S) \quad , \quad \delta B = \varepsilon (B,S), \quad \delta C = \varepsilon (C,S) \tag{8.3}$$

where S is the generating function of the canonical transformation and ε is an infinitesimal constant. Now, the Poisson brackets standing on the right of (8.3) are simply functions of the dynamical variables, just as A, B, C; etc. are. The quantum analogs of these Poisson brackets must simply be the quantum analogs of the functions which they represent. Substituting (8.3) into (8.2) we obtain the quantum expression

$$\delta F = \varepsilon \left[(A,S) BC \ldots + A(B,S) C \ldots + AB (C,S) \ldots + \ldots \right] \tag{8.4}$$

But we must also have

$$\delta F = \varepsilon (F,S) \tag{8.5}$$

Hence, since ε is arbitrary

$$(F,S) = (A,S) BC \ldots + A(B,S) C \ldots + AB(C,S) \ldots + \ldots \tag{8.6}$$

By means of equation (8.6) we may now obtain a special expression for Poisson brackets in the quantum theory. Restricting our attention to the case of three observables A; B, C, we may evidently write, as a special case of (8.1) and (8.6)

$$(AB,C) = (AC,B) + A(B,C). \tag{8.7}$$

Inverting the Poisson brackets, we may also write (8.7) in the form

$$(C,AB) = (CA,B) + A(C,B). \tag{8.8}$$

Now consider a Poisson bracket of the form (AB, CD). With the use of

eqs. (8.7) and (8.8) this Poisson bracket may be expanded in two different ways, thus :

$$(AB,CD) = (A,CD)B + A(B,CD)$$

$$= [(A,C)D + C(A,D)]B + A[(B,C)D + C(B,D)] \quad (8.9)$$

and

$$(AB,CD) = (AB,C)D + C(AB,D)$$

$$= [(A,C)B + A(B,C)]D + C[(A,D)B + A(B,D)] \quad (8.10)$$

Substracting equation (8.9) from equation (8.10) we obtain

$$0 = (A,C)(BD - DB) - (AC - CA)(B,D) \quad (8.11)$$

or

$$(A,C)[B,D] = [A,C](B,D). \quad (8.12)$$

Since equation (8.12) must hold for all observables A, B, C, D, we may evidently infer

$$(A,B) = \lambda [A,B] \quad (8.13)$$

where A and B are any two proper real operators and λ is some universal constant. Since A and B are real $[A, B]$ is pure imaginary, and λ must be pure imaginary in order to make $(A, B,)$ real. From the Correspondence Principle argument which resulted in equations (6.17) and (6.19), we may evidently make the identification $\lambda = (i\hbar)^{-1}$, so that

$$(A,B) = \frac{1}{i\hbar}[A,B]. \quad (8.14)$$

One may readily show that commutator brackets satisfy equations completely analogous to the equations (1.38-40) satisfied by Poisson brackets, namely

$$[F_1, F_2] = -[F_2, F_1], \quad [F_1, F_2+F_3] = [F_1, F_2] + [F_1, F_3], \tag{8.15}$$

$$[F_1, F_2+F_3] = [F_1, F_2]F_3 + F_2[F_1, F_3], \tag{8.16}$$

$$[F_1, [F_2, F_3]] + [F_2, [F_3, F_1]] + [F_3, [F_1, F_2]] = 0. \tag{8.17}$$

Problem XLIII : Verify eqs (8.15-17)

The following relation, involving an anticommutator bracket, is also sometimes useful

$$\begin{aligned}[F_1, F_2F_3] &= F_1F_2F_3 - F_2F_3F_1 \\ &= F_1F_2F_3 + F_2F_1F_3 - F_2F_1F_3 - F_2F_3F_1 \\ &= \{F_1, F_2\}F_3 - F_2\{F_3, F_1\}.\end{aligned} \tag{8.18}$$

If equation (8.16) is written in the form

$$[F_1, F_2F_3] = [F_1, F_2]F_3 - F_2[F_3, F_1], \tag{8.19}$$

it is seen to possess a certain analogy with (8.18)

The dynamical behavior of quantum operators is contained in the equation (see (6.20)).

$$\dot{F} = (F, H) = \frac{1}{i\hbar}[F, H] \tag{8.20}$$

where F is an arbitrary observable and H is the operator corresponding to the Hamiltonian function of the system in question. Strictly speaking, equation (8.20) is correct only when F possesses no explicit dependence on the time. The more general equation, allowing for the possibility of such an explicit time dependence is the analog of equation (1.64) :

$$\dot{F} = \frac{1}{i\hbar}[F,H] + \frac{\partial F}{\partial t} \tag{8.21}$$

An equation which is more directly usable in the prediction of the results of physical measurements is the following,

$$\overline{\dot{F}} = \frac{d}{dt}\langle |F| \rangle = \langle |\dot{F}| \rangle = \frac{1}{i\hbar}\langle |[F,H]| \rangle + \langle \left|\frac{\partial F}{\partial t}\right| \rangle, \tag{8.22}$$

which gives the time rate of change of the expectation value of any observable in a quantum state $|\ \rangle$. In obtaining the expression on the right hand side of (8.22) we made use of the fact that

$$\frac{d}{dt}|\ \rangle = 0\ ,\quad \frac{d}{dt}\langle\ | = 0\ . \tag{8.23}$$

That is, the state vector of the system is assumed to be constant in time. This corresponds to a point of view in which the "state" of the system is determined once and for all by an all-embracing symbol $|\ \rangle$. This state can be fixed at some one instant of time by a measurement of the various dynamical variables (to the extent that these variables can be simultaneously measured under the fundamental limitation imposed by equation (7.149). The expectation values of these observables will then be determined for all subsequent times by equation (8.22). The changes in the observed configurations, etc. of the system are taken into account through the time variation of the quantum operators.

This notion of "state" differs somewhat from the classical notion of "state" which actually refers to the coordinates and momenta at any instant and changes with time. We shall see presently,

however, that by introducing a suitable unitary transformation we can obtain a quantum representation of "state" which is conceptually closer to the classical notion. In this representation, on the other hand, the quantum equations are less analogous to the classical equations.

The fundamental equations (8.21) are often most easily solved in terms of some particular representation. As we have seen in the preceding chapter, representations are defined with respect to complete sets of commuting proper real operators. The present system of dynamical equations is seen to be not so convenient for the purposes of representation theory. For, unless we happen to choose the commuting operators to be constants of the motion, they will change with time, and a representation which refers to the values of these operators at one instant of time will be of little practical use at a later time. The way out of this difficulty is as follows.

Consider an operator $U(t'', t')$ which is defined by the integral equation

$$U(t'', t') = 1 + \frac{1}{i\hbar}\int_{t'}^{t''} U(t, t')H(t)\,dt \qquad (8.24)$$

Here a possible explicit time dependence of the Hamiltonian operator is indicated by writing it in the form H(t). Setting t' = t" in (8.24), we obtain

$$U(t', t') = 1. \qquad (8.25)$$

It is evident, from the differentiation of eq. (8.24), that the operator U (t", t') also satisfies the differential equation

$$i\hbar \frac{\partial}{\partial t''} U(t'', t') = U(t'', t')H(t'') \qquad (8.26)$$

The boundary condition (8.25) is not contained in the differential equation, but must be added to it.

The operator U (t",t') is readily shown to be unitary. The proof proceeds by induction. First, take the complex conjugate (or

Hermitian adjoint) of eq. (8.26) :

$$-i\hbar \frac{\partial}{\partial t''} U^*(t'',t') = H(t'') U^*(t'',t') \tag{8.27}$$

Using eqs. (8.26, 27) we may write

$$\frac{\partial}{\partial t''}\left[U(t'',t')U^*(t'',t')\right] = \frac{1}{i\hbar} U(t'',t')\left[H(t'')-H(t'')\right]U^*(t'',t') = 0, \tag{8.28}$$

Hence $U(t'', t')\,U^*(t'', t')$ is constant and independent of t". But $U(t', t')\,U^*(t', t')$ for all t'. Hence

$$U(t'',t')\,U^*(t'',t') = 1 \quad \text{for all } t', t'' \tag{8.29}$$

Similarly, from

$$\frac{\partial}{\partial t''}\left[U^*(t'',t)U(t'',t')\right] = \frac{1}{i\hbar}\left[U^*(t'',t)U(t'',t'), H(t'')\right] \tag{8.30}$$

we may infer

$$U^*(t'',t')U(t'',t') = U^*(t',t')U(t',t') + \frac{(t''-t')}{i\hbar}\left[U^*(t',t)U(t',t'), H(t')\right]_{t'} \tag{8.31}$$
$$= 1 \qquad \text{for all } t', t''$$

Now, take the complex conjugate of equation (8.24):

$$U^*(t'',t') = 1 - \frac{1}{i\hbar}\int_{t'}^{t''} H(t)\,U^*(t,t')\,dt \tag{8.32}$$

Introduce the operator

$$V(t'',t,t''') \equiv U^*(t'',t')\,U^*(t,t''')$$
$$= U^*(t,t''') - \frac{1}{i\hbar}\int_{t'}^{t''} H(t)\,V(t,t',t''')\,dt \tag{8.33}$$

Differentiating eq. (8.33) with respect to t" we obtain

$$-i\hbar \frac{\partial}{\partial t''} V(t'',t,t''') = H(t'')\,V(t'',t,t'''). \tag{8.34}$$

$V(t'', t', t''')$ is seen to satisfy the same differential equation as $U^*(t'', t''')$. Moreover, setting $t'' = t'$, we have the boundary condition $V(t', t', t''') = U^*(t', t''')$. Thus $V(t'', t', t''')$ is identical with $U^*(t'', t''')$ for all values of t'', and is independent of t'. We therefore have the relation

$$U^*(t'', t')\ U^*(t', t''') = U^*(t'', t''') \tag{8.35}$$

Setting $t''' = t''$ in (8.35) we obtain

$$U^*(t'', t')\, U^*(t', t'') = U^*(t'', t'') = 1 \tag{8.36}$$

from which we may infer

$$U^*(t'', t') = U(t', t''). \tag{8.37}$$

Now, let $|\rangle$ be the state vector of the physical system in question. $|\rangle$, as we have remarked, is independent of the time t. Let us introduce a time-dependent state-vector $|t\rangle$ defined by

$$|t\rangle = U(t, t_0)|\rangle, \quad |\rangle = U^*(t, t_0)|t\rangle \tag{8.38}$$

where t_0 is some chosen fixed point of time. We may now write the expectation value of any observable F at the time t in the form

$$\overline{F} = \langle|F|\rangle = \langle t|U(t, t_0)FU^*(t, t_0)|t\rangle = \langle t|F_s|t\rangle \tag{8.39}$$

where

$$F_s = U(t, t_0)\, F\, U^*(t, t_0). \tag{8.40}$$

Differentiating equation (8.40) with respect to t, and using eqs. (8.21), (8.26) and (8.27), we obtain

$$\dot{F}_s = \dot{U}(t,t_0)F\,U^*(t,t_0) + U(t,t_0)\dot{F}U^*(t,t_0) + U(t,t_0)F\dot{U}^*(t,t_0)$$
$$= \frac{1}{i\hbar}U(t,t_0)\left\{H(t)F + [F,H(t)] - H(t)F\right\}U^*(t,t_0)$$
$$+ U(t,t_0)\frac{\partial F}{\partial t}U^*(t,t_0)$$
$$= \frac{\partial F_s}{\partial t}. \qquad (8.41)$$

That is, the time variation of the operators F_s is due solely to the explicit time dependence which they may have. If F is a function of the canonical coordinates and momenta alone, having no explicit dependence on the time, then

$$\dot{F}_s = 0. \qquad (8.42)$$

The description of the system in terms of the state-vectors $|\rangle$ and operators F is called the Heisenberg representation We have now introduced a new representation, known as the Schrödinger representation, in which the state-vectors $|t\rangle$ and operators F_s give the description of the system. The operators corresponding to the canonical coordinates and momenta (and all functions of them) are seen to be constant in the Schrödinger representation, and the task of describing the temporal behavior of the system is borne by the state-vectors $|t\rangle$. These Schrödinger state-vectors satisfy the following differential equation

$$i\hbar\frac{d}{dt}|t\rangle = i\hbar\dot{U}(t,t_0)|\rangle = U(t,t_0)H(t)|\rangle$$
$$= U(t,t_0)H(t)U^*(t,t_0)|t\rangle = H_s(t)|t\rangle \qquad (8.43)$$

Equation (8.43) is known as the Schrödinger equation.

Since the transformation (8.38) is unitary the state-vector $|t\rangle$ will be normalized if $|\ \rangle$ is. The Heisenberg and Schrödinger state-vectors evidently coincide at the time t_0. So also do the Heisenberg and Schrödinger operators :

$$|t_0\rangle = |\ \rangle \tag{8.44}$$

$$F_S = F(t_0) \tag{8.45}$$

If the Hamiltonian function is time independent, then

$$U_S(t'',t') = e^{\frac{1}{i\hbar}H(t''-t_0)}\, e^{\frac{1}{i\hbar}H(t_0-t')}$$
$$= e^{\frac{1}{i\hbar}H_0(t''-t')} = U(t'',t')$$

In the case in which the Hamiltonian function has no explicit time dependence, and is therefore constant, the integral equation (8.24) can readily be solved. Its solution is

$$U(t'',t') = e^{\frac{1}{i\hbar}H(t''-t')} \tag{8.46}$$

We have further,

$$H_S = H \tag{8.47}$$

and

$$F(t) = U^*(t,t_0)\, F_S\, U(t,t_0)$$
$$= e^{-\frac{1}{i\hbar}H(t-t_0)}\, F(t_0)\, e^{\frac{1}{i\hbar}H(t-t_0)} \tag{8.48}$$

Using eq. (4.) of the Appendix, we may expand (8.48) in the form

$$F(t) = F(t_0) + \frac{t-t_0}{i\hbar}[F(t_0),H] + \frac{1}{2!}\left(\frac{t-t_0}{i\hbar}\right)^2[[F(t_0),H],H] + \tag{8.49}$$

This is completely analogous to the classical expansion (cf. (1.63)

$$F(t) = F(t_0) + (t-t_0)(F(t_0),H) + \frac{1}{2!}(t-t_0)^2((F(t_0),H),H) + \ldots \quad (8.50)$$

Having introduced the Schrödinger representation, we are now at liberty to introduce various convenient vector bases. In order to define a particular basis we must choose a complete set of commuting proper real operators. Since we are working in the Schrödinger representation we know that these operators will remain constant in time, and our basic vectors will remain fixed. One of the most useful representations is that in which the basic vectors are chosen as eigenvectors of the generalized coordinates ρ^i out of which the Lagrangian function of the system is constructed. From the commutation relations (see (6.26), (8.14))

$$[\rho^i, \rho^j] = 0, \quad [p_i, p_j] = 0, \quad [\rho^i, p_j] = i\hbar\delta^i_j \quad (8.51)$$

one may readily see that the ρ^i form a complete set. First of all, they commute with each other; secondly, they are Hermitian, or real, since the corresponding classical quantities are real ; thirdly, any operator which commutes with all of them must be a function of them, since if it is a function of any of the ρ's it will not commute with the ρ's. Finally, we shall see presently, by carrying out an explicit construction, that the ρ's are proper, i.e. their eigenvectors form a complete set.

It is almost intuitively evident why the representation defined by the ρ's should be a particularly useful one. One of the most direct experiments one can make on a system is an observation of its configuration (for example, an observation of successive positions of a particle in a cloud-chamber ; an observation of the position of a beam of particles, as in a Stern-Gerlach experiment). The configuration is described in terms of the ρ's. In a typical experiment, using the configuration point of view, we would try to determine the probability of a system's being found in a certain

region of its configuration space.

Actual experience tells us something about the eigenvalues of the p's. In measurements of a particular coordinate of a system we can in general expect to obtain results ranging over a continuum of values. Let us therefore begin by supposing that each of the p^i possesses a continuum of eigenvalues p'^i ranging from $-\infty$ to $+\infty$. The first task that we then have is to deal with the difficulty of summing over the p'^i. We employ a limiting process. Let $f(p')$ be an arbitrary function of the p'^i. Then we write

$$\sum_{p'} f(p') \Delta\omega = \int f(p') d\omega \qquad (8.52)$$

where $\Delta\omega$ is a small constant volume in p'-space. That is, we divide the whole of p'-space up into cells of volume $\Delta\omega$ and understand a Σ summation over the p'^i to be a summation over a set of points with one point to a cell. Equations of the form (8.52) are then always to be understood as holding in the limit of vanishing $\Delta\omega$. That is, we employ an implicit limiting convention.

We must next make sure that the limiting procedure which we have set up is an invariant one. For example, suppose we introduce a different set of coordinates. Since the new coordinates are simply functions of the old, we known that any measurement of the old coordinates gives us immediately a measurement of the new, and vice versa. All coordinate systems are diagonal in any coordinate representation. Moreover, we have seen in chapter 1, that the Lagrangian form of the equations of motion of any dynamical system remains invariant under any coordinate transformation. Therefore we are at liberty to begin from any coordinate system we please ; all are equally valid. A summation such as (8.52) can be expressed equally well in terms of any coordinate representation. We must, however, make sure that the cell-volume $\Delta\omega$ can be chosen independently of the coordinate system.

So far we have not considered the question of defining volumes in p'-space. In order to give an invariant prescription for

$\Delta\omega$ we must take a closer look at the physics of the system under onsideration. We shall now make a very important statement : <u>All physical systems which possess classical analogs are describable in terms of Lagrangian functions of the form</u>

$$L = \frac{1}{2}\mu\, g_{ij}\dot{\rho}_i\dot{\rho}_j + A_i\dot{\rho}_i - V \qquad (8.53)$$

<u>where the g_{ij}, A_i, and V are functions of the ρ's, A_i and V are possibly also functions of the time t, and where the matrix (g_{ij}) is symmetric and non-singular</u>. This, of course, is a very drastic statement, as it greatly restricts the possible forms which the Lagrangian functions of dynamical systems can have. Nevertheless, it appears to be satisfied in nature, and much of our subsequent discussion depends upon it. If, at any time in the future, a classically describable dynamical system is found for which this statement is not true, then a great deal of revision and generalization of the work which follows will be necessary.

From a study of typical, well-known systems of the type (8.53), for example, a particle of mass m constrained to move on the surface of a sphere of radius R, for which

$$L = \frac{1}{2}m\left(R^2\dot{\theta}^2 + R^2\sin^2\theta\,\dot{\varphi}^2\right), \qquad (8.54)$$

we may infer that the ρ's are coordinates in an n-dimensional Riemannian space of which g_{ij} is the metric tensor. n may actually be non-denumerably infinite (e.g. in field theories) and the constant μ may not always be directly interpretable as a mass. Nevertheless we shall continue to speak as if ρ-space were a perfectly ordinary Riemannian space, and we shall refer to μ as the "mass constant".

With the use of the metric tensor we may now give an invariant characterization of the cell-volume $\Delta\omega$. We should perhaps best begin with a brief review of the theory of Riemannian

spaces. Let us for the time being drop the primes on the eigenvalues of the operators φ^i. Primes will be used to denote coordinate transformations. Let $\delta_\mu \varphi^i$, $\mu = 1 \ldots m$ be a set of m infinitesimal displacements at an arbitrary point φ. The m-dimensional volume $\delta_m \omega$ enclosed by the parallelepiped defined by these displacements is defined to be

$$\delta_m \omega^2 = | \delta_\mu \varphi^i \varphi_{ij} \delta_\nu \varphi^j | \tag{8.55}$$

That is, $\delta_m \omega$ is equal to the square root of the determinant formed from the m^2 scalar products of the $\delta_\mu \varphi^i$ with one another. Two cases are of special interest, namely $m = 1$ and $m = n$. In the case $m = 1$, $\delta_1 \omega$ is called the invariant length element and is often denoted by δs:

$$\delta s^2 = \varphi_{ij} \delta \varphi^i \delta \varphi^j \tag{8.56}$$

In the case $m = n$, $\delta_n \omega$ is called the invariant volume element and is denoted simply by $\delta \omega$:

$$\delta \omega^2 = |\varphi_{ij}| \, |\delta_\mu \varphi^i|^2 \tag{8.57}$$

Problem XLI : In passing from eq. (8.55) to eq. (8.57), use is made of the fact that the determinant of the product of two or more square matrices is equal to the product of their determinants. Prove this. Hint : Use the alternating symbol $\varepsilon_{i_1 \ldots i_n}$.

If the displacements are taken along the coordinate mesh, as in an integral, so that

$$\delta_\mu \varphi^i = \delta_\mu^i \, d\varphi^i \quad , \quad \text{N.S.,} \tag{8.58}$$

then

$$\delta\omega^2 = g\,(dp^1 \ldots dp^n)^2. \tag{8.59}$$

where g denotes the determinant of the metric tensor. The infinitesimal volume element $d\omega$ in an integration is evidently

$$d\omega = g^{1/2}\, dp^1 \ldots dp^n. \tag{8.60}$$

$d\omega$ is an invariant.

An invariant is a quantity whose value remains unchanged under a coordinate transformation. It is often referred to as a scalar. A covariant vector A_i is a set of quantities which transform like the derivatives of a scalar. If φ is a scalar, we have, using commas as in chapter 3 to denote differentiation with respect to coordinates

$$\varphi'_{,i} = \frac{\partial\varphi}{\partial p'^i} = \frac{\partial\varphi}{\partial p^r}\frac{\partial p^r}{\partial p'^i} = \frac{\partial p^s}{\partial p'^i}\varphi_{,s} \tag{8.61}$$

Hence

$$A'_i = \frac{\partial p^s}{\partial p'^i} A_s. \tag{8.62}$$

A contravariant vector B^i is a set of quantities which transform like the components of an infinitesimal displacement. We have

$$\delta p'^i = \frac{\partial p'^i}{\partial p^s}\delta p^s \tag{8.63}$$

Hence

$$B'^i = \frac{\partial p'^i}{\partial p^j} B^j \tag{8.64}$$

A tensor is a quantity which can always be expressed as a sum of products of vectors, e.g.

$$T_{ij\ldots}^{\ kl\ldots} = \sum_{\mu} A_{\mu i} B_{\mu j} \ldots C_{\mu}^{k} D_{\mu}^{l} \ldots \qquad (8.65)$$

The transformation laws of tensors can be inferred directly from (8.62) and (8.64). A tensor with lower indices only is called a covariant tensor, a tensor with upper indices only is called a contravariant tensor, and a tensor with both upper and lower indices is called a mixed tensor. The Kronecker delta is a mixed tensor. Thus

$$\delta_i^{\,i} = \frac{\partial \rho'^{\,i}}{\partial \rho^{k}} \frac{\partial \rho^{k}}{\partial \rho'^{\,i}} = \frac{\partial \rho'^{\,i}}{\partial \rho^{k}} \frac{\partial \rho^{e}}{\partial \rho'^{\,i}} \delta_e^{\,k} \qquad (8.66)$$

From (8.62), (8.64) and (8.66) one may readily see that the scalar product of a covariant vector and a contravariant vector is a scalar.

$$A'_i B'^{\,i} = A_i B^i \qquad (8.67)$$

Since expression (8.56) is postulated to be an invariant, the metric tensor ρ_{ij} must be a covariant vector. A contravariant form of it, ρ^{ij}, can be introduced through the definition.

$$\rho_{ik}\rho^{kj} = \rho^{jk}\rho_{ki} = \delta_i^{\,j}. \qquad (8.68)$$

The matrix (ρ^{ij}) is evidently the inverse of (ρ_{ij}). With the use of the two forms of the metric tensor, indices are "raised" and "lowered" according to

$$A^{ij}{}_{k} = \rho^{je}\rho_{km} A_e^{\,im}, \text{ etc.} \qquad (8.69)$$

Evidently $\rho^{i}{}_{j} = \delta_j^{\,i} = \delta^{i}{}_{j}$.

The transformation laws (8.62), (8.64), etc. must be taken to be evaluated at the point at which the vectors (or tensors) involved are defined. A vector may be defined at only one point (e.g. a displacement), or it may be defined over many points as a function of position. In the latter case, the concept of the covariant

derivative is very useful. The covariant derivative of a covariant vector A_i is defined by

$$A_{i;j} \equiv A_{i,j} - \Gamma^{I}_{ij}{}^{k} A_k \tag{8.70}$$

and the covariant derivative of a contravariant vector B^i is defined by

$$B^{i}_{;j} \equiv B^{i}_{,j} + \Gamma^{II}_{kj}{}^{i} B^k . \tag{8.71}$$

The coefficients $\Gamma^{I}_{ij}{}^{k}$, $\Gamma^{II}_{kj}{}^{i}$ are determined by a number of geometrical requirements which we shall now proceed to enumerate.

First of all, covariant derivatives themselves must transform like tensors (or vectors). This requirement can be shown to impose a unique transformation law on the $\Gamma^{I}_{ij}{}^{k}$, $\Gamma^{II}_{kj}{}^{i}$. We shall postpone the derivation of this law for the present, however. Secondly it shall be required that the process of contraction (i.e. placing two indices equal and summing over them) give the same result whether it is carried out before or after a covariant differentiation. The covariant differentiation laws of tensors can be obtained by applying eqs. (8.70, 71) to (8.65) via the distributive law. Now, it will be remembered that the ordinary derivative of a scalar is already a vector quantity (see (8.61)). It is assumed, therefore, that the covariant derivative of a scalar is simply its ordinary derivative. We may now infer, for arbitrary A_i and B^i,

$$(A_i B^i)_{;j} = (A_i B^i)_{,j} = A_{i;j} B^i + A_i B^i{}_{;j}$$

$$= (A_{i,j} - \Gamma^{I}_{ij}{}^{k} A_k) B^i + A_i (B^i{}_{,j} + \Gamma^{II}_{kj}{}^{i} B^k)$$

$$= (A_i B^i)_{,j} - (\Gamma^{I}_{ij}{}^{k} - \Gamma^{II}_{ij}{}^{k}) A_k B^i$$

(8.72)

Since A_i and B^i are arbitrary we must have

$$\Gamma^{I}_{ij\,k} = \Gamma^{II}_{ij\,k} = \left\{ \begin{matrix} k \\ ij \end{matrix} \right\}. \tag{8.73}$$

$\left\{ \begin{matrix} k \\ ij \end{matrix} \right\}$ is called a Christoffel symbol of the first kind.

Indices induced by repeated covariant differentiation do not in general commute. However, there is one case in which we shall require them to commute, namely, when two of them are applied to a scalar. In this way we insure that our invariant formalism is as close as possible to the formalism with which we are familiar in the case of Cartesian coordinate systems and will reduce to it when g_{ij} reduces to the Cartesian metric δ_{ij}. We now have

$$\begin{aligned} 0 = \varphi_{,ij} - \varphi_{,ji} &= (\varphi_{,i})_{,j} - (\varphi_{,j})_{,i} \\ &= \varphi_{,ij} - \left\{ \begin{matrix} k \\ ij \end{matrix} \right\} \varphi_{,k} - \varphi_{,i} + \left\{ \begin{matrix} k \\ ji \end{matrix} \right\} \varphi_{,k} \\ &= \left[\left\{ \begin{matrix} k \\ ji \end{matrix} \right\} - \left\{ \begin{matrix} k \\ ij \end{matrix} \right\} \right] \varphi_{,k} \end{aligned} \tag{8.74}$$

Since $\varphi_{,k}$ is arbitrary we must have

$$\left\{ \begin{matrix} k \\ ij \end{matrix} \right\} = \left\{ \begin{matrix} k \\ ji \end{matrix} \right\} \tag{8.75}$$

In the case of covariant differentiation of a vector we may now write

$$\begin{aligned} A_{k\,ij} - A_{k\,ji} &= (A_{k,i})_{,j} - \left\{ \begin{matrix} \ell \\ kj \end{matrix} \right\} A_{\ell,i} - \left\{ \begin{matrix} \ell \\ ij \end{matrix} \right\} A_{k,\ell} \\ &\quad - (A_{k,j})_{,i} + \left\{ \begin{matrix} \ell \\ ki \end{matrix} \right\} A_{\ell,j} + \left\{ \begin{matrix} \ell \\ ji \end{matrix} \right\} A_{k,\ell} \\ &= \left(A_{k,i} - \left\{ \begin{matrix} \ell \\ ki \end{matrix} \right\} A_\ell \right)_{,j} - \left\{ \begin{matrix} \ell \\ kj \end{matrix} \right\} \left(A_{\ell,i} - \left\{ \begin{matrix} m \\ \ell i \end{matrix} \right\} A_m \right) \\ &\quad - \left(A_{k,j} - \left\{ \begin{matrix} \ell \\ kj \end{matrix} \right\} A_\ell \right)_{,i} + \left\{ \begin{matrix} \ell \\ ki \end{matrix} \right\} \left(A_{\ell,j} - \left\{ \begin{matrix} m \\ \ell j \end{matrix} \right\} A_m \right) \end{aligned}$$

$$= - R_{ijk}{}^{\ell} A_{\ell}, \tag{8.76}$$

where

$$R_{ijk}{}^{\ell} = \left\{ \begin{matrix} \ell \\ ik \end{matrix} \right\}_{,j} - \left\{ \begin{matrix} \ell \\ jk \end{matrix} \right\}_{,i} + \left\{ \begin{matrix} \ell \\ im \end{matrix} \right\} \left\{ \begin{matrix} m \\ jk \end{matrix} \right\} - \left\{ \begin{matrix} \ell \\ jm \end{matrix} \right\} \left\{ \begin{matrix} m \\ ik \end{matrix} \right\}. \tag{8.77}$$

$R_{ijk}{}^{\ell}$ is called the curvature tensor.

As a final requirement on the Christoffel symbols it is to be postulated that the process of raising or lowering indices give the same result no matter whether it is performed before or after a covariant differentiation. Thus

$$g_{ik} A^{k}{}_{,j} = A_{ij} = (g_{ik} A^{k})_{,j} = g_{ik,j} A^{k} + g_{ik} A^{k}{}_{,j} \tag{8.78}$$

This implies that the covariant derivative of the metric tensor vanishes. Introducing the Christoffel symbol of the second kind,

$$[ij, k] \equiv g_{k\ell} \left\{ \begin{matrix} \ell \\ ij \end{matrix} \right\} \tag{8.79}$$

we may write*

$$0 = g_{ij}{}^{,k} = g_{ij,k} - \left\{ \begin{matrix} \ell \\ ik \end{matrix} \right\} g_{\ell j} - \left\{ \begin{matrix} \ell \\ jk \end{matrix} \right\} g_{i\ell} \tag{8.80}$$

or

$$g_{ij,k} = [ik, j] + [jk, i] \tag{8.81}$$

Therefore, finally,

---

* Since $\delta^{i}_{j,k} = \delta^{i}_{j,k} + \left\{ \begin{matrix} i \\ \ell k \end{matrix} \right\} \delta^{\ell}_{j} - \left\{ \begin{matrix} \ell \\ jk \end{matrix} \right\} \delta^{i}_{\ell} = \left\{ \begin{matrix} i \\ jk \end{matrix} \right\} - \left\{ \begin{matrix} i \\ jk \end{matrix} \right\} = 0,$
we also have $g^{ij}{}_{,k} = 0.$

$$\frac{1}{2}(g_{ik,j}+g_{jk,i}-g_{ij,k}) = \frac{1}{2}\big([ij,k]+[kj,i] +[ji,k]+[ki,j] -[ik,j]-[jk,i]\big)$$

$$= [ij,k]. \qquad (8.82)$$

From (8.82) we may now derive the transformation laws of the Christoffel symbols. We have

$$g'_{ij,k} = \frac{\partial g'_{ij}}{\partial \rho'^k} = \frac{\partial \rho^e}{\partial \rho'^k}\frac{\partial}{\partial \rho^e}\left(\frac{\partial \rho^m}{\partial \rho'^i}\frac{\partial \rho^n}{\partial \rho'^j} g_{mn}\right)$$

$$= \frac{\partial \rho^e}{\partial \rho'^k}\frac{\partial \rho^m}{\partial \rho'^i}\frac{\partial \rho^n}{\partial \rho'^j} g_{mn,e} + \left(\frac{\partial \rho^m}{\partial \rho'^j}\frac{\partial^2 \rho^n}{\partial \rho'^i \partial \rho'^k} + \frac{\partial \rho^m}{\partial \rho'^i}\frac{\partial^2 \rho^n}{\partial \rho'^j \partial \rho'^k}\right) g_{mn} \qquad (8.83)$$

and hence

$$[ij,k]' = \frac{\partial \rho^e}{\partial \rho'^k}\frac{\partial \rho^m}{\partial \rho'^i}\frac{\partial \rho^n}{\partial \rho'^j}[mn,e] + \frac{\partial \rho^m}{\partial \rho'^k}\frac{\partial^2 \rho^n}{\partial \rho'^i \partial \rho'^j} g_{mn} \qquad (8.84)$$

and

$$\left\{ {k \atop ij} \right\}' = \frac{\partial \rho'^k}{\partial \rho^e}\frac{\partial \rho^m}{\partial \rho'^i}\frac{\partial \rho^n}{\partial \rho'^j}\left\{ {e \atop mn} \right\} + \frac{\partial \rho'^k}{\partial \rho^e}\frac{\partial^2 \rho^e}{\partial \rho'^i \partial \rho'^j}. \qquad (8.85)$$

Using (8.85), one may verify directly that expressions (8.70, 71) transform like tensors.

Problem XLV : Verify this.

Returning now to the study of invariant volume elements, let us first observe that the determinant of the metric tensor is expressible, in terms of the alternating symbol $\varepsilon^{i_1 \cdots i_m}$, in the form

$$g = \frac{1}{m!} \varepsilon^{i_1 \cdots i_m} \varepsilon^{j_1 \cdots j_m} g_{i_1 j_1} \cdots g_{i_m j_m} . \tag{8.86}$$

The alternating symbol has the following transformation law :

$$\varepsilon^{i_1 \cdots i_m} = \left| \frac{\partial \rho'}{\partial \rho} \right|^{-1} \frac{\partial \rho'^{i_1}}{\partial \rho^{j_1}} \cdots \frac{\partial \rho'^{i_m}}{\partial \rho^{j_m}} \varepsilon^{j_1 \cdots j_m} . \tag{8.87}$$

Hence the metric determinant transforms according to

$$g' = \left| \frac{\partial \rho'}{\partial \rho} \right|^{-2} g . \tag{8.88}$$

The elements of the contravariant metric tensor may be expressed in terms of those of the covariant metric tensor by the following equation

$$g^{ij} = g^{-1} \varepsilon^{i_1 \cdots i_{j-1} i \, i_{j+1} \cdots i_m} g_{1 i_1} \cdots g_{j-1 \, i_{j-1}} g_{j+1 \, i_{j+1}} \cdots g_{m i_m} \tag{8.89}$$

Using (8.89) one may obtain the following useful relation involving the Christoffel symbol :

$$\begin{aligned} \left(\log g \right)_{,i} &= g^{-1} g_{,i} = g^{-1} \left(\varepsilon^{i_1 \cdots i_m} g_{1 i_1} \cdots g_{m i_m} \right)_{,i} \\ &= g^{-1} \varepsilon^{i_1 \cdots i_m} \left(g_{1 i_1,i} g_{2 i_2} \cdots g_{m i_m} + \cdots + g_{1 i_1} g_{2 i_2} \cdots g_{m i_m,i} \right) \\ &= g^{i_1 1} g_{1 i_1,i} + \cdots + g^{i_m m} g_{m i_m,i} = g^{jk} g_{jk,i} \\ &= g^{jk} \left(g_{ik,j} + g_{jk,i} - g_{ij,k} \right) = 2 g^{jk} [ij,k] = 2 \left\{ {j \atop ij} \right\} \end{aligned} \tag{8.90}$$

Let us now revert to the practice of putting primes on the ρ's to indicate that the Riemannian space in which we are working is the space of eigenvalues of the complete set of operators ρ^i. Let ρ' and ρ'' denote two points of this space. It will be convenient to introduce the generalized delta-function $\delta(\rho', \rho'')$, defined by

$$\int f(\rho'') \delta(\rho', \rho'') \, d\omega'' = f(\rho') \quad \text{for all functions } f \tag{8.91}$$

The representation of the invariant volume element in the form (8.60) or, in the present case,

$$d\omega'' = \rho^{1/2}(\rho'') \, d\rho''^1 \ldots d\rho''^n \tag{8.92}$$

enables one to put the generalized delta-function into the following analytic form involving the ordinary delta-function

$$\delta(\rho', \rho'') = \rho^{-1/2}(\rho'') \delta(\rho' - \rho'') = \rho^{-1/2}(\rho') \delta(\rho' - \rho'') \tag{8.93}$$

Using this form one may derive some useful identities. With the aid of eqs. (1.19) and (1.23) of the Appendix, we may write

$$(\rho'^i - \rho''^i) \delta(\rho', \rho'') = \rho^{-1/2}(\rho') (\rho'^i - \rho''^i) \delta(\rho' - \rho'') = 0 \tag{8.94}$$

$$(\rho'^i - \rho''^i) \frac{\partial}{\partial \rho''^j} \delta(\rho', \rho'') = \rho^{-1/2}(\rho'') (\rho'^i - \rho''^i) \delta(\rho' - \rho'')$$

$$= -\rho^{-1/2}(\rho'') \delta^i_j \delta(\rho' - \rho'') = -\delta^i_j \delta(\rho', \rho'') \tag{8.95}$$

$$\frac{\partial}{\partial \rho'^i} \delta(\rho'', \rho'') = \frac{\partial}{\partial \rho'^i} [\rho^{-1/2}(\rho') \delta(\rho' - \rho'')]$$

$$= -\frac{1}{2} \rho^{-3/2}(\rho') \rho_{,i}(\rho') \delta(\rho' - \rho'') + \rho^{-1/2}(\rho') \frac{\partial}{\partial \rho'^i} \delta(\rho' - \rho'')$$

$$= -\frac{1}{2}\rho^{-1}(\rho')\,g_{ij}(\rho')\,\delta(\rho',\rho'') - \rho^{-\frac{1}{2}}(\rho')\frac{\partial}{\partial\rho''^{i}}\,\delta(\rho'-\rho'')$$

$$= -\frac{\partial}{\partial\rho''^{i}}\,\delta(\rho',\rho'') - \left\{ {j \atop ij} \right\}\delta(\rho',\rho'') \qquad (8.96)$$

The contracted Christoffel symbol appearing in (8.96) may be evaluated at either ρ' or ρ''.

The matrix elements of the operators ρ^{i} in the coordinate representation are given by

$$\langle \rho' / \rho^{i} / \rho'' \rangle = \rho'^{i}\,\delta_{\rho'\rho''} = \rho''^{i}\,\delta_{\rho'\rho''}\,, \qquad (8.97)$$

the eigenvectors $/\rho'\rangle$ being normalized so that

$$\langle \rho' / \rho'' \rangle = \delta_{\rho'\rho''}\,. \qquad (8.98)$$

The Kronecker delta may be expressed in terms of the generalized delta function via the limiting process implied by equation (8.52). Let f be an arbitrary function of the ρ's. Then

$$\int f(\rho'')\,\delta(\rho',\rho'')\,d\omega'' = f(\rho') = \sum_{\rho''} f(\rho'')\,\delta_{\rho'\rho''}$$

$$= \sum_{\rho''} \frac{f(\rho'')}{\Delta\omega}\,\delta_{\rho'\rho''}\,\Delta\omega = \int \frac{f(\rho'')}{\Delta\omega}\,\delta_{\rho'\rho''}\,d\omega'' \qquad (8.99)$$

Since f is arbitrary, we may infer

$$\delta_{\rho'\rho''} = \Delta\omega\,\delta(\rho',\rho'') \qquad (8.100)$$

The matrix representation of the operators p_i can now be obtained from the commutation relations (8.51). In matrix form the third of these equations becomes

$$i\hbar\,\delta^{i}_{j}\,\delta_{\rho'\rho''} = \langle \rho' | [\rho^{i}, p_{j}] | \rho'' \rangle$$

$$= \sum_{\rho'''} \left\{ \langle \rho' | \rho^{i} | \rho''' \rangle \langle \rho''' | p_{j} | \rho'' \rangle - \langle \rho' | p_{j} | \rho''' \rangle \langle \rho''' | \rho^{i} | \rho'' \rangle \right\}$$

$$= (\rho'^{i} - \rho''^{i}) \langle \rho' | p_{j} | \rho'' \rangle . \qquad (8.101)$$

Using equations (8.94), (8.95) and (8.100) we see that $\langle \rho' | p_{j} | \rho'' \rangle$ must have the form

$$\langle \rho' | p_{j} | \rho'' \rangle = -i\hbar \frac{\partial}{\partial \rho'^{i}} \delta_{\rho'\rho''} + F_{j}\,\delta_{\rho'\rho''} \qquad (8.102)$$

where F_{j} is an arbitrary function of the ρ's which may be evaluated at either ρ' or ρ''.

Expression (8.102) is the most general solution of equation (8.101). This solution must also satisfy the second of the commutation relations (8/51). This will impose certain conditions on the functions F_{j}. Introducing the functions

$$\overline{F}_{i} = F_{i} + i\hbar \left\{ \begin{matrix} \delta \\ ij \end{matrix} \right\} \qquad (8.103)$$

we may write

$$0 = \langle \rho' | [p_{i}, p_{j}] | \rho'' \rangle$$

$$= \sum_{\rho'''} \left\{ \langle \rho' | p_{i} | \rho''' \rangle \langle \rho''' | p_{j} | \rho'' \rangle - \langle \rho' | p_{j} | \rho''' \rangle \langle \rho''' | p_{i} | \rho'' \rangle \right\}$$

$$= \frac{1}{\Delta\omega} \int \left\{ \left[-i\hbar \frac{\partial}{\partial \rho'^{i}} \delta_{\rho'\rho'''} + F_{i}(\rho''') \delta_{\rho'\rho'''} \right] \left[-i\hbar \frac{\partial}{\partial \rho'''^{j}} \delta_{\rho'''\rho''} + F_{j}(\rho'') \delta_{\rho'''\rho''} \right] \right.$$

$$\left. - \left[-i\hbar \frac{\partial}{\partial \rho'^{j}} \delta_{\rho'\rho'''} + F_{j}(\rho''') \delta_{\rho'\rho'''} \right] \left[-i\hbar \frac{\partial}{\partial \rho'''^{i}} \delta_{\rho'''\rho''} + F_{i}(\rho''') \delta_{\rho'''\rho''} \right] \right\} d\omega'''$$

$$= \Delta\omega\, \rho^{-1/2}(\rho') \int \Big\{ \Big[-ik\hbar \frac{\partial}{\partial \rho'^i} \delta(\rho' - \rho''') + \bar{F}_i(\rho''') \delta(\rho' - \rho''') \Big] \Big[ik\hbar \frac{\partial}{\partial \rho'''^j} \delta(\rho''' - \rho'') + \bar{F}_j(\rho''') \delta(\rho''' - \rho'') \Big]$$

$$- \Big[-ik\hbar \frac{\partial}{\partial \rho''^j} \delta(\rho' - \rho''') + \bar{F}_j(\rho''') \delta(\rho' - \rho''') \Big] \Big[-ik\hbar \frac{\partial}{\partial \rho'''^i} \delta(\rho''' - \rho'') + \bar{F}_i(\rho''') \delta(\rho''' - \rho'') \Big] \Big\} d\rho'''$$

$$= \Delta\omega\, \rho^{-1/2}(\rho') \Big\{ \hbar^2 \frac{\partial^2}{\partial \rho''^j \partial \rho''^i} \delta(\rho' - \rho'') + ik\hbar \frac{\partial}{\partial \rho''^j} \bar{F}_i(\rho'') \delta(\rho' - \rho'') + ik\hbar \bar{F}_i(\rho'') \frac{\partial}{\partial \rho''^j} \delta(\rho' - \rho'')$$

$$- ik\hbar\, \bar{F}_j(\rho'') \frac{\partial}{\partial \rho''^i} \delta(\rho' - \rho'') + \bar{F}_i(\rho'') \bar{F}_j(\rho'') \delta(\rho' - \rho'')$$

$$- \hbar^2 \frac{\partial^2}{\partial \rho'''^i \partial \rho''^j} \delta(\rho' - \rho''') - ik\hbar \frac{\partial}{\partial \rho'''^i} \bar{F}_j(\rho''') \delta(\rho' - \rho''') - ik\hbar\, \bar{F}_j(\rho''') \frac{\partial}{\partial \rho'''^i} \delta(\rho' - \rho''')$$

$$+ ik\hbar\, \bar{F}_i(\rho'') \frac{\partial}{\partial \rho'^j} \delta(\rho' - \rho'') - \bar{F}_j(\rho''') \bar{F}_i(\rho''') \delta(\rho' - \rho''') \Big\}$$

$$= ik\hbar \Big[\frac{\partial}{\partial \rho'^j} \bar{F}_i(\rho') - \frac{\partial}{\partial \rho'^i} \bar{F}_j(\rho') \Big] \delta_{\rho'\rho''}, \tag{8.104}$$

which implies that the $\bar{F}_i$ are of the form

$$\bar{F}_i(\rho') = \frac{\partial}{\partial \rho'^i} F(\rho'), \tag{8.105}$$

where $F(\rho')$ is some function independent of the index i. Since the Christoffel symbol $\{ {}^{\alpha}_{\beta\gamma} \}$ is also of this form, being in fact equal to $1/2\ (\log \rho)_{,i}$ we may write

$$\langle \rho' | p_i | \rho'' \rangle = -ik\hbar \frac{\partial}{\partial \rho'^i} \delta_{\rho'\rho''} + \frac{\partial F(\rho')}{\partial \rho'^i} \delta_{\rho'\rho''} \tag{8.106}$$

where

$$F = \bar{F} - ik\hbar \log \rho^{1/2}. \tag{8.107}$$

Owing to the fact that the ρ's are Hermitian operators the function F is not entirely arbitrary. The Hermitian condition may be expressed in the form

$$\langle \rho' | p_i | \rho'' \rangle = \langle \rho'' | p_i | \rho' \rangle^*, \tag{8.108}$$

or

$$-i\hbar \frac{\partial}{\partial \rho'^i} \delta_{\rho'\rho''} + \frac{\partial F(\rho')}{\partial \rho'^i} \delta_{\rho'\rho''} = i\hbar \frac{\partial}{\partial \rho''^i} \delta_{\rho''\rho'} + \frac{\partial F^*(\rho'')}{\partial \rho''^i} \delta_{\rho''\rho'}$$

$$= -i\hbar \frac{\partial}{\partial \rho''^i} \delta_{\rho'\rho''} - i\hbar \left\{ \begin{matrix} j \\ ij \end{matrix} \right\} \delta_{\rho'\rho''} + \frac{\partial F^*(\rho')}{\partial \rho'^i} \delta_{\rho'\rho''} \tag{8.109}$$

which implies

$$F^* = F + i\hbar \log \rho^{1/2}. \tag{8.110}$$

F must therefore have the form

$$F = R - \frac{1}{2} i\hbar \log \rho^{1/2} \tag{8.111}$$

where R is a real function of the ρ's.

The function R may be removed from the scene by a trivial unitary transformation. If we introduce a new set of basic eigenvectors $|\rho'\rangle^+$ which are connected with the old eigenvectors by the relation

$$|\rho'\rangle^+ = e^{-\frac{1}{i\hbar} R(\rho')} |\rho'\rangle, \tag{8.112}$$

the matrix elements of the p_i then reduce to

$$^+\langle \rho' | p_i | \rho'' \rangle^+ = \sum_{\rho''', \rho'''} {}^+\langle \rho' | \rho''' \rangle \langle \rho''' | p_i | \rho''' \rangle \langle \rho''' | \rho'' \rangle^+$$

$$= e^{-\frac{1}{i\hbar} R(\rho')} \langle \rho' | p_i | \rho'' \rangle e^{\frac{1}{i\hbar} R(\rho'')}$$

$$= -i\hbar \frac{\partial}{\partial \rho'^i}\left[e^{-\frac{i}{\hbar}R(\rho')}\delta_{\rho'\rho''}e^{\frac{i}{\hbar}R(\rho'')}\right] - \frac{\partial R(\rho')}{\partial \rho'^i}\left[e^{-\frac{i}{\hbar}R(\rho')}\delta_{\rho'\rho''}e^{\frac{i}{\hbar}R(\rho'')}\right]$$

$$+\frac{\partial E(\rho')}{\partial \rho'^i}\left[e^{-\frac{i}{\hbar}R(\rho')}\delta_{\rho'\rho''}e^{\frac{i}{\hbar}R(\rho'')}\right]$$

$$= -i\hbar \frac{\partial}{\partial \rho'^i}\delta_{\rho'\rho''} - \frac{1}{2}i\hbar\left\{\begin{matrix} i \\ ij \end{matrix}\right\}\delta_{\rho'\rho''}. \qquad (8.113)$$

The transformation (8.112) is called a phase transformation. We shall always suppose that this transformation has already been carried out, so that the matrix elements of the p_i are given immediately by (8.113).

Now let $|t\rangle$ be the Schrödinger state-vector of the system whose Lagrangian function is (8.93). The coordinate representation which has been constructed above has been defined with respect to a set of constant Schrödinger operators, and to be completely consistent with a previous notation we should put the subscript ζ on all the ρ's, ρ''s, ρ'''s, etc. However, the ζ has been omitted for brevity, as we are always at liberty to think of the ρ's and p's as the corresponding Heisenberg operators taken at the time t_0, when the Heisenberg and Schrödinger representations coincide. The probability that a measurement of the system at the time t will find the configuration of the system located in a certain volume Ω of ρ-space is

$$\sum_{\rho' \text{ in } \Omega} |\langle \rho' | t\rangle|^2 = \sum_{\rho' \text{ in } \Omega} \frac{|\langle \rho' | t\rangle|^2}{\Delta\omega}\Delta\omega = \int_\Omega |\psi(\rho', t)|^2 \, d\omega' \qquad (8.114)$$

where

$$\psi(\rho', t) \equiv \Delta\omega^{-\frac{1}{2}}\langle \rho' | t\rangle. \qquad (8.115)$$

$\psi(\rho',t)$ is called the wave function of the system. $|\psi(\rho',t)|^2 d\omega'$ evidently gives the probability that the system will be found in the element $d\omega'$ of configuration space at the point ρ' at the time t'. Since $d\omega'$ is an invariant, and probability is an invariant, $\psi(\rho',t)$ must be an invariant or scalar. Since the state-vector $|t\rangle$ is normalized we have the following condition on the wave function :

$$1 = \langle t|t\rangle = \sum_{\rho'} \langle t|\rho'\rangle\langle\rho'|t\rangle = \sum_{\rho'} \frac{|\langle\rho'|t\rangle|^2}{\Delta\omega}\Delta\omega$$
$$= \int |\psi(\rho',t)|^2 d\omega'. \qquad (8.116)$$

That is, the probability that the system will be found somewhere is unity.

It is often more convenient to deal with the wave-functions $\psi(\rho',t')$ than with the probability amplitudes $\langle\rho'|t\rangle$. We shall therefore be interested in determining the wave function which results when the state-vector is acted upon by an arbitrary linear operator. In particular, we shall write

$$p'\psi(\rho',t') \equiv \Delta\omega^{-\frac{1}{2}}\langle\rho'|p'|t\rangle$$
$$= \Delta\omega^{-\frac{1}{2}} \sum_{\rho''} \langle\rho'|p'|\rho''\rangle\langle\rho''|t\rangle$$
$$= \sum_{\rho''} p'' \delta_{\rho'\rho''}\psi(\rho'',t) = p''\psi(\rho',t) \qquad (8.117)$$

and

$$p_i\psi(\rho',t) \equiv \Delta\omega^{-\frac{1}{2}}\langle\rho'|p_i|t\rangle$$
$$= \Delta\omega^{-\frac{1}{2}} \sum_{\rho''} \langle\rho'|p_i|\rho''\rangle\langle\rho''|t\rangle$$

$$= \frac{1}{\Delta\omega}\int\left(-i\hbar\frac{\partial}{\partial\rho''_i}\delta_{\rho'\rho''} - \frac{1}{2}i\hbar\left\{\frac{\partial}{\partial\rho'_i},\frac{\partial}{\partial\rho''_i}\right\}\delta_{\rho'\rho''}\right)\psi(\rho'',t)\,d\omega''$$

$$= \int\left[i\hbar\frac{\partial}{\partial\rho'_i}\delta(\rho'-\rho'') - \frac{1}{2}i\hbar\left\{\frac{\partial}{\partial\rho'_i},\frac{\partial}{\partial\rho''_i}\right\}\delta(\rho'-\rho'')\right]\psi(\rho'',t)\,d_n\rho''$$

$$= -i\hbar\frac{\partial}{\partial\rho'_i}\psi(\rho',t) - \frac{1}{2}i\hbar\left\{\frac{\partial}{\partial\rho'_i},f(\rho')\right\}\psi(\rho',t). \tag{8.118}$$

From equations (8.117, 118) we see that we may write symbolically

$$\underline{\rho}^i = \rho'^i \tag{8.119}$$

$$\underline{p}_i = -i\hbar\frac{\partial}{\partial\rho'^i} - \frac{1}{2}i\hbar\left\{\frac{\partial}{\partial\rho'^i},f(\rho')\right\} \tag{8.120}$$

More generally

$$F(\underline{\rho}^i,\underline{p}_i) = F\left(\rho'^i, -i\hbar\frac{\partial}{\partial\rho'^i} - \frac{1}{2}i\hbar\left\{\frac{\partial}{\partial\rho'^i},f(\rho')\right\}\right). \tag{8.121}$$

From now on, when we work with the wave functions instead of the probability amplitudes, we shall drop the primes used to designate particular eigenvalues of the coordinate operators, and we shall drop the bars underlining the letters on the left sides of equations (8.119, 121). The use of primes will be reserved to designate general coordinate transformations.

In chapter I we have seen that in classical mechanics, for every transformation of the generalized coordinates of the form

$$\rho'^i = \rho'^i(\rho), \quad \left|\frac{\partial\rho'}{\partial\rho}\right| \neq 0, \tag{8.122}$$

there exists a corresponding transformation of the conjugate momenta, of the form

$$p'_i = \frac{\partial \rho^j}{\partial \rho'^i} p_j \,, \qquad (8.123)$$

which preserves the canonical nature of the ρ's and p's. The p_i are evidently the components of a covariant vector.

Equations (8.122) and (8.123) together define what was called, in chapter I, a point transformation. Point transformations may also be defined in quantum mechanics. The only possible difficulty in extending the notion to quantum mechanics would be an ambiguity in ordering non-commutiong factors in the quantum analog of a classical expression. Since the ρ's all commute with one another there is clearly non ambiguity in writing down the quantum analog of eq. (8.122). There is likewise actually no ambiguity in obtaining the quantum analog of eq. (8.123). For the only problem here is that of correctly symmetrizing the right hand side of (8.123) so as to make it Hermitian. One may easily convince oneself that all methods of symmetrization lead to the same result, namely,

$$p'_i = \frac{1}{2}\left\{\frac{\partial \rho^j}{\partial \rho'^i}, p_j\right\} \qquad (8.124)$$

For example, one might expand $\partial \rho^j / \partial \rho'^i$ in a power series in the ρ's. The operator p_j could then be inserted between the ρ's in any symmetrical fashion in each term of the series. The result of commuting p_j symmetrically to the left and to the right through the ρ's would be to produce two terms of order $\hbar$ which cancel each other, leaving simply the expression (8.124).

That equation (8.124) gives the correct transformation law for the momentum operators may also be verified by making explicit use of expressions (8.119, 120). We have

$$p'_i = \frac{1}{2}\frac{\partial \rho^j}{\partial \rho'^i} p_j + \frac{1}{2} p_j \frac{\partial \rho^j}{\partial \rho'^i} = \frac{\partial \rho^j}{\partial \rho'^i} p_j + \frac{1}{2}\left[p_j, \frac{\partial \rho^j}{\partial \rho'^i}\right]$$

$$= \frac{\partial \rho^{\alpha}}{\partial \rho'^{i}} \left[-i\hbar \frac{\partial}{\partial \rho^{\alpha}} - \frac{1}{2} i\hbar \left\{ {k \atop j\alpha} \right\} \right] - \frac{1}{2} i\hbar \frac{\partial}{\partial \rho^{\alpha}} \frac{\partial \rho^{\alpha}}{\partial \rho'^{i}}$$

$$= -i\hbar \frac{\partial}{\partial \rho'^{i}} - \frac{1}{2} i\hbar \left\{ {\alpha \atop ij} \right\}' \qquad (8.125)$$

where

$$\left\{ {i \atop jk} \right\}' = \frac{\partial \rho'^{i}}{\partial \rho^{\alpha}} \left\{ {\alpha \atop jk} \right\} + \frac{\partial \rho'^{i}}{\partial \rho^{\alpha}} \frac{\partial^2 \rho^{\alpha}}{\partial \rho'^{j} \partial \rho'^{k}} \qquad (8.126)$$

Referring to (8.85), we see that eq. (8.126) is just the transformation law for the constructed Christoffel symbol.

It is now evident that there exists an isomorphism between the group of point transformations in classical mechanics and a corresponding subgroup of the group of all unitary transformations in quantum mechanics. The group property ensures that each point transformation has an inverse. It is instructive to display explicitly the inverse of eq. (8.124). We need first the lemma

$$\{A, \{B, C\}\} = ABC + ACB + BCA + CBA$$
$$- ABC + BAC + CAB + CBA$$
$$+ BCA - BAC - CAB + ACB,$$

$$= \{\{A, B\}, C\} + [B, [C, A]] \qquad (8.127)$$

We may then write

$$\frac{1}{2} \left\{ \frac{\partial \rho'^{i}}{\partial \rho^{\alpha}}, p_{\alpha} \right\} + \frac{1}{2} \left\{ \frac{\partial \rho'^{i}}{\partial \rho^{\alpha}}, \left\{ \frac{\partial \rho^{\alpha}}{\partial \rho'^{k}}, p_{k} \right\} \right\}$$
$$= \frac{1}{4} \left\{ \left\{ \frac{\partial \rho'^{i}}{\partial \rho^{\alpha}}, \frac{\partial \rho^{\alpha}}{\partial \rho'^{k}} \right\}, p_{k} \right\} + \frac{1}{4} \left[\frac{\partial \rho'^{i}}{\partial \rho^{\alpha}}, \left[p_{k}, \frac{\partial \rho^{\alpha}}{\partial \rho'^{k}} \right] \right]$$
$$= \frac{1}{2} \{ \delta_{ik}, p_{k} \} - \frac{1}{4} i\hbar \left[\frac{\partial \rho'^{i}}{\partial \rho^{\beta}}, \frac{\partial^2 \rho^{\beta}}{\partial \rho'^{k} \partial \rho'^{k}} \right] = p_{i} \qquad (8.128)$$

which shows that the inverse transformation has the same form as (8.124).

The unitary representations of the point transformation group may be obtained by determining the infinitesimal generators of the group. An infinitesimal point transformation may be expressed in the form

$$\rho'^i = \rho^i + \epsilon \Lambda^i(\rho), \tag{8.129}$$

$$p'_i = p_i - \frac{1}{2}\epsilon \left\{ \frac{\partial \Lambda^r(\rho)}{\partial \rho^i} p_r \right\}, \tag{8.130}$$

where ϵ is an infinitesimal constant and Λ^i is a function of the ρ's. The result of carrying out the point transformation (8.129, 130) on an arbitrary function F of the ρ's and p's is

$$F' = F + \frac{\epsilon}{i\hbar}[F, S] \tag{8.131}$$

where

$$S = \frac{1}{2}\left\{ \Lambda^i(\rho), p_i \right\}. \tag{8.132}$$

S is the generating function of the infinitesimal point transformation.

The subgroup of unitary transformations in quantum mechanics which corresponds isomorphically to the group of all proper point transformations in classical mechanics is given by the set of all unitary operators $e^{\tau/i\hbar S}$, where S has the form (8.132) and τ is an arbitrary parameter. Each set of functions Λ^i defines a one-parameter subgroup of the point transformation group.

The canonical momenta for the Lagrangian function (8.53) are

$$p_i = \frac{\partial L}{\partial \dot{\rho}^i} = \hbar \rho_{ij} \dot{\rho}^j + A_i . \tag{8.133}$$

Solving (8.133) for the $\dot{\rho}_i$, we obtain

$$\dot{\rho}^i = \frac{1}{\mu} g^{ij}(p_j - A_j) \tag{8.134}$$

and therefore the Hamiltonian function becomes

$$\begin{aligned} H &= p_i \dot{\rho}^i - L \\ &= \mu g_{ij} \dot{\rho}^i \dot{\rho}^j + A_i \dot{\rho}^i - \frac{1}{2}\mu g_{ij} \dot{\rho}^i \dot{\rho}^j - A_i \dot{\rho}^i + V \\ &= \frac{1}{2}\mu g_{ij} \dot{\rho}^i \dot{\rho}^j + V \\ &= \frac{1}{2\mu} g^{ij}(p_i - A_i)(p_j - A_j) + V \end{aligned} \tag{8.135}$$

In writing the quantum analog of this we are faced with the problem of properly ordering non-commuting factors. The problem presents itself here in a more serious form than that which we dealt with in obtaining the quantum analog of (8.123), because we have terms quadratic as well as linear in the momenta. The linear terms we of course symmetrize by the anticommutator method, as before, so the problem reduces to that of finding a prescription for determining the quantum analog of $g^{ij} p_i p_j$. We must require of this prescription that it be invariant, i.e. that it be independent of the coordinate system in which we carry it out.

To get an idea of how to proceed, let us write down the Schrödinger equation for the wave function. We have

$$\begin{aligned} i\hbar \frac{\partial \psi}{\partial t} &= i\hbar \Delta\omega^{-1/2} \frac{\partial}{\partial t} \langle \rho' | t \rangle = i\hbar \Delta\omega^{-1/2} \langle \rho' | \left(\frac{d}{dt} | t\rangle\right) \\ &= \Delta\omega^{-1/2} \langle \rho' | H | t \rangle = H(\rho, p, t)\, \psi \end{aligned} \tag{8.136}$$

where the ρ's and p's in the last expression are the operators

(8.119, 120). Let us now recall that ψ is a scalar. The left hand side of eq. (8.136) is therefore a scalar, and hence its right hand side, $H\psi$, must also be a scalar. In particular

$$(g^{ij} p_i p_j)_{\substack{\text{quantum}\\ \text{analog}}} \psi \tag{8.137}$$

must be a scalar. Now, it will be seen from (8.120) that the operators p_i are essentially differential operators. It is therefore reasonable to assume that expression (8.137) is actually equal to the invariant expression

$$\begin{aligned} -\hbar^2 g^{ij} \psi_{;ij} &= -\hbar^2 g^{ij} \left[(\psi_{,i})_{,j} - \left\{ {k \atop ij} \right\} \psi_{,k} \right] \\ &= -\hbar^2 g^{ij} \left[\psi_{,ij} - \left\{ {k \atop ij} \right\} \psi_{,k} \right], \end{aligned} \tag{8.138}$$

so that

$$(g^{ij} p_i p_j)_{\substack{\text{quantum}\\ \text{analog}}} = -\hbar^2 g^{ij} \left[\frac{\partial^2}{\partial \rho^i \partial \rho^j} - \left\{ {k \atop ij} \right\} \frac{\partial}{\partial \rho^k} \right]. \tag{8.139}$$

We must now ask ourselves how we can arrange the factors in the expression $g^{ij} p_i p_j$ so as to produce the operator (8.139). The simplest method of symmetrizing the expression $g^{ij} p_i p_j$ so as to make it Hermitian is to write it in the form

$$\begin{aligned} p_i g^{ij} p_j = &\left(-i\hbar \frac{\partial}{\partial \rho^i} - \frac{1}{2} i\hbar \left\{ {k \atop ik} \right\} \right) g^{ij} \left(-i\hbar \frac{\partial}{\partial \rho^j} - \frac{1}{2} i\hbar \left\{ {l \atop jl} \right\} \right) \\ &- \hbar^2 \left[g^{ij}{}_{,i} \left(\frac{\partial}{\partial \rho^j} + \frac{1}{2} \left\{ {l \atop jl} \right\} \right) \right. \\ &\left. + g^{ij} \left(\frac{\partial^2}{\partial \rho^i \partial \rho^j} + \frac{1}{2} \left\{ {l \atop jl} \right\}_{,i} + \left\{ {k \atop ik} \right\} \frac{\partial}{\partial \rho^j} + \frac{1}{4} \left\{ {k \atop ik} \right\} \left\{ {l \atop jl} \right\} \right) \right] \end{aligned} \tag{8.140}$$

Now, using (8.90) and $g^{kj}{}_{,i} = -g^{kl} g^{jm} g_{lm,i}$, we find

$$g^{ij}\left\{\begin{matrix} k \\ ij \end{matrix}\right\} = g^{ij} g^{kl} [ij, l] = \frac{1}{2} g^{ij} g^{kl} (g_{il,j} + g_{jl,i} - g_{ij,l})$$
$$= -g^{kj}{}_{,j} - g^{kl} \left\{\begin{matrix} j \\ lj \end{matrix}\right\}. \qquad (8.141)$$

Hence, substracting (8.140) from (8.139) we obtain

$$-\hbar^2 g^{ij}\left[\frac{\partial^2}{\partial\rho^i \partial\rho^j} - \left\{\begin{matrix} k \\ ij \end{matrix}\right\}\frac{\partial}{\partial\rho^k}\right] - p_i g^{ij} p_j$$
$$= -\hbar^2\left[g^{ij}\frac{\partial^2}{\partial\rho^i\partial\rho^j} + g^{kj}{}_{,j}\frac{\partial}{\partial\rho^k} + g^{kl}\left\{\begin{matrix} j \\ lj \end{matrix}\right\}\frac{\partial}{\partial\rho^k}\right.$$
$$- g^{ij}{}_{,i}\frac{\partial}{\partial\rho^j} + \frac{1}{2}\left\{\begin{matrix} l \\ jl \end{matrix}\right\}\left(g^{kl}\left\{\begin{matrix} j \\ kl \end{matrix}\right\} + g^{jl}\left\{\begin{matrix} k \\ lk \end{matrix}\right\}\right)$$
$$\left. - g^{ij}\left(\frac{\partial^2}{\partial\rho^i\partial\rho^j} + \frac{1}{2}\left\{\begin{matrix} l \\ jl \end{matrix}\right\}_{,i} + \left\{\begin{matrix} k \\ ik \end{matrix}\right\}\frac{\partial}{\partial\rho^j} + \frac{1}{4}\left\{\begin{matrix} k \\ ik \end{matrix}\right\}\left\{\begin{matrix} l \\ jl \end{matrix}\right\}\right)\right] \qquad (8.142)$$
$$= \hbar^2 Q$$

where

$$Q = \frac{1}{2} g^{ij}\left[\left\{\begin{matrix} k \\ ik \end{matrix}\right\}_{,j} - \left\{\begin{matrix} k \\ ij \end{matrix}\right\}\left\{\begin{matrix} l \\ kl \end{matrix}\right\} - \frac{1}{2}\left\{\begin{matrix} k \\ ik \end{matrix}\right\}\left\{\begin{matrix} l \\ jl \end{matrix}\right\}\right] \qquad (8.143)$$

Expressions (8.139) and (8.140) are therefore, in general, not equal to each other. Furthermore, it is actually impossible to find any method of symmetrizing the expression $g^{ij} p_i p_j$ so as to make it equal to (8.139). The symmetrical quantities will always differ from (8.139) by a quantity of order $\hbar^2$ which is a function of the ρ's alone. Here wa see a case in which the quantum analog of a classical quantity is not obtained simply by turning the ρ's and p's into operators and symmetrizing them in a suitable fashion. We must also add a quantity proportional to $\hbar^2$. This is obviously permissable, however, even on the basis of the Correspondence Principle, since in the classical limit $\hbar \to 0$, the quantum analog

still reduces to the classical quantity.

We now have, for the quantum Hamiltonian operator

$$H = \frac{1}{2\mu}(p_i - A_i)\rho^{ij}(p_j - A_j) + \frac{\hbar^2}{2\mu}Q + V,$$

$$= \frac{1}{2\mu}(p_i\rho^{ij}p_j + \hbar^2 Q) - \frac{1}{2\mu}\{A^i, p_i\} + \frac{1}{2\mu}A^iA_i + V \quad (8.144)$$

Since

$$\frac{1}{2}\{A^i, p_i\}\psi = (A^i p_i + \frac{1}{2}[p_i, A_i])\psi$$

$$= -i\hbar\left(A^i\frac{\partial}{\partial\rho^i} + \frac{1}{2}A^i{}_{,j}\rho^{ij} + \frac{1}{2}A^i{}_{,i}\right)\psi$$

$$= -i\hbar\left(A^i\frac{\partial}{\partial\rho^i} + \frac{1}{2}A^i{}_{,i}\right)\psi, \quad (8.145)$$

the Schrödinger equation (8.136) takes the explicit form

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2\mu}\rho^{ij}\psi_{,ij} + \frac{i\hbar}{\mu}\left(A^i\psi_{,i} + \frac{1}{2}A^i{}_{,i}\psi\right)$$

$$+\left(\frac{1}{2\mu}A^iA_i + V\right)\psi. \quad (8.146)$$

A few final remarks should be made about the above formalism. We began by supposing that the eigenvalues of each ρ^i form a continuum ranging from $-\infty$ to $+\infty$. For many important systems this is, however, not the case, e.g. a particle constrained to move on the surface of a sphere, a free particle viewed in spherical coordinates, etc.. The Schrödinger equation (8.146) is nevertheless still valid in these cases. The only change in the above formalism necessitated by such situations is the imposition of cyclic conditions on the wave functions.

The above formalism is unsymmetrical in the coordinates and momenta. Such an asymmetry does not exist in the transformation theory of classical mechanics, where one deals only with Poisson brackets in which coordinates and momenta are on a completely equal footing. Here we have attached special importance to a small subgroup of the group of all canonical transformations, namely the point-transformation group. This subgroup is singled out by the special forms of the Lagrangian function (8.53) and Hamiltonian function (8.125) with which we are concerned. To be sure, in the classical theory, one can make a canonical transformations of a completely arbitrary type, thus destroying the special Hamiltonian form (8.125). The question that imposes itself is therefore the following : given a Hamiltonian function, what conditions must it satisfy in order that it be canonically transformable into a Hamiltonian which is quadratic in the momenta ? The answer to this question is not easy. In practice we are never bothered by this problem since we always start from the Lagrangian function of a classical system, which, as we have stated, always has the form (8.53)

---

* Problem XLVI : What the conditions under which an arbitrary function of the p's and q's is canonically transformable into a function which is quadratic (though inhomogeneously so) in the momenta.

A last remark should be made on an important application of equation (7.149) of the preceding chapter. From the commutation relations (8.51) we find

$$\Delta p^i \Delta p_i \geq \frac{1}{2} \langle | -i [p^i, p_i] | \rangle = \frac{\hbar^2}{2} , \quad i \neq j \qquad (8.147)$$

That is, the product of the root mean square deviation in the measu-

rement of a coordinate with the root mean square deviation in the measurement of the conjugate momentum is always greater than or equal to $\hbar/2$. This is known as Heisenberg's Uncertainty Principle. It is quite analogous to the uncertainty principle; involving time and energy, expressed by equation (5.112).

9.- SPECIAL SYSTEMS.

a) The free particle in Cartesian coordinates.

Let the coordinates be x_i, i = 1, 2, 3. Using the position vector $r = (x_1, x_2, x_3)$ for short, we may express the Lagrangian function in the form

$$L = \frac{1}{2} m \dot{r}^2 \tag{9.1}$$

where m is the particle mass. The conjugate momenta p_1, p_2, p_3 form the components of a vector p given by

$$p = \frac{\partial L}{\partial \dot{r}} = m \dot{r} . \tag{9.2}$$

The Hamiltonian function is

$$H = p \cdot \dot{r} - L = \frac{1}{2m} p^2 , \tag{9.3}$$

and in the quantum theory we have the commutation relations

$$[x_i, x_j] = 0, \quad [p_i, p_j] = 0, \quad [x_i, p_j] = i \hbar \delta_{ij} \tag{9.4}$$

Let us first work in the Heisenberg representation. Since

$$[p, H] = 0 \tag{9.5}$$

the momentum is a constant of the motion. (This means, incidentally that the operator p is unchanged on passage to the Schrödinger representation). The position vector r, however, is not a constant of the motion. If we indicate its time dependence by writing it in the form $r(t)$, we have, using eq. (8.49),

$$r(t) = r(t') + \frac{t-t'}{i\hbar}[r(t), H] + \frac{1}{2!}\left(\frac{t-t'}{i\hbar}\right)^2 [[r(t), H], H] + \ldots$$
$$= r(t') + \frac{t-t'}{m} p, \qquad (9.6)$$

which can also be obtained by integrating equation (9.2) directly. The commutation law for the coordinates at two different times then immediately follows :

$$[x_i(t), x_j(t')] = [x_i(t'), x_j(t')] + \frac{t-t'}{m}[p_i, x_j(t')]$$
$$= -\delta_{ij}\frac{i\hbar}{m}(t-t') \qquad (9.7)$$

Using the uncertainty relation (7.149) we see that the root mean square deviations of two successive measurements of the position of a free particle are related by

$$\Delta x_i(t)\,\Delta x_i(t') \geq \frac{\hbar}{2m}|t-t'|, \text{ N.S.} \qquad (9.8)$$

Evidently the two measurements interefere with each other more and more as the time between them increases. The time dependence of the interference is linear. Also, the greater the mass of the particle the less is the interefference. In the classical limit of macroscopic bodies and large masses the interefference becomes completely unobservable.

Let us now go over the Schrödinger representation. A very useful set of basis vectors for many problems are the eigenvectors $|p'\rangle$ of the momentum operator :

$$p|p'\rangle = p'|p'\rangle. \qquad (9.9)$$

A question immediately arises as to the distribution of the eigenvalues p'. In the preceding chapter we saw how one could construct a consis-

tent theory by assuming, on the basis of analogy with classical theory, that the eigenvalues of the coordinates range continuously from $-\infty$ to $+\infty$ (except for special cases which require various cyclic conditions). We should expect similarly that momentum eigenvalues will range from $-\infty$ to ∞. The validity of this expectation can readily be verified. Let k be an arbitrary vector, and consider the state vector

$$e^{ik\cdot r}|p'\rangle. \tag{9.10}$$

Multiplying this vector on the left by the operator p, we obtain

$$\begin{aligned} p e^{ik\cdot r}|p'\rangle &= [p, e^{ik\cdot r}]|p'\rangle + e^{ik\cdot r} p|p'\rangle \\ &= \hbar k e^{ik\cdot r}|p'\rangle + e^{ik\cdot r} p'|p'\rangle \\ &= (p' + \hbar k)(e^{ik\cdot r}|p'\rangle) \end{aligned} \tag{9.11}$$

Hence the vector (9.10) is an eigenvector of p corresponding to the eigenvalue $p' + \hbar k$. Since k is arbitrary it is evident that if p possesses a single eigenvalue it possesses a continuum of eigenvalues covering the totality of momentum space.

The wave-functions corresponding to the momentum eigenvectors are readily obtained. Since the Cartesian metric is δ_{ij}, the differential form (8.120) of the momentum operator reduces to

$$p = -i\hbar\nabla.$$

If $\psi_{p'}(r)$ denotes the wave function corresponding to the vector $|p'\rangle$, we have

$$p\psi_{p'}(r) = -i\hbar\nabla\psi_{p'}(r) = p'\psi_{p'}(r), \tag{9.13}$$

of which the solution is

$$\psi_{p'}(r) = A\, e^{i/\hbar\, p'\cdot r}. \tag{9.14}$$

The constant A may be determined by the condition that the state vectors $|p'\rangle$ be normalized. Referring to equations (8.52) and (8.115), we may write

$$\begin{aligned}\delta_{p'p''} = \langle p''|p'\rangle &= \sum_r \langle p''|r'\rangle\langle r'|p'\rangle \\ &= \frac{1}{\Delta\omega}\int \langle p''|r'\rangle\langle r'|p'\rangle\, d_3r' = \int \psi^*_{p''}(r)\,\psi_{p'}(r)\, d^3r \\ &= |A|^2 \int e^{i/\hbar\,(p'-p'')\cdot r}\, d_3r \\ &= |A|^2 (2\pi)^3 \delta(k'-k'')\end{aligned} \tag{9.15}$$

where

$$\hbar k' = p', \quad \hbar k'' = p'' \tag{9.16}$$

But

$$\delta_{p'p''} = \delta_{k'k''} = \frac{\delta(k'-k'')}{\delta(k'-k')}. \tag{9.17}$$

Hence, at a fixed instant of time we may choose

$$A = (2\pi)^{-3/2}[\delta(k'-k')]^{-1/2}. \tag{9.18}$$

Or, using the notational convention of equation (2.31) of the Appendix, we have

$$A = \frac{1}{\sqrt{L^3}} \tag{9.19}$$

where L^3 is the "volume of space". Thus finally

$$\langle r'|p'\rangle = \Delta\omega^{1/2}\psi_{p'}(r') = \sqrt{\frac{\Delta\omega}{L^3}}\, e^{\frac{i}{\hbar}p'\cdot r'}$$

$$= \Delta\omega^{1/2}(2\pi)^{-3/2}\left[\delta(k'-k')\right]^{-1/2} e^{ik'\cdot r'} \qquad (9.20)$$

That the vectors $|p'\rangle$ form a complete set follows directly from Fourier transform theory. The condition for completeness may be expressed concisely in the form

$$\sum_{p'} |p'\rangle\langle p'| = 1. \qquad (9.21)$$

Using the limiting convention implied by eqs. (2.10) and (2.35) of the Appendix, the above sum may be replaced by an integral. We have

$$\sum_{p'} \langle r''|p'\rangle\langle p'|r'\rangle = \frac{\Delta\omega}{(2\pi)^3}\sum_{k'} \frac{e^{ik'\cdot(r''-r')}}{\delta(k'-k')}$$

$$= \frac{\Delta\omega}{(2\pi)^3}\int e^{ik'(r''-r')}\, d_3k' = \Delta\omega\,\delta(r''-r')$$

$$= \delta_{r''r'} = \langle r''|1|r'\rangle \qquad (9.22)$$

Since $\langle r''|$ and $|r'\rangle$ are arbitrary eigenvectors of r, eq. (9.21) follows. The momentum operator p is therefore proper.

Let us now pause to observe that there is an apparent contradiction hidden in the foregoing formalism. Consider the eigenvector equation (9.9). Since p is supposed to be a real operator, the complex conjugate of this equation is simply

$$\langle p'|p = \langle p'|p'. \qquad (9.23)$$

Hence we are led to write equations of the form

$$i\hbar\,\delta_{ij} = \langle p'|i\hbar\,\delta_{ij}|p'\rangle = \langle p'|[x_i, p_j]|p'\rangle$$
$$= \langle p'|x_i p_j|p'\rangle - \langle p'|p_j x_i|p'\rangle$$
$$= (p'_j - p'_j)\langle p'|x_i|p'\rangle = 0, \qquad (9.24)$$

which is plainly nonsense.

The existence of such contradictions indicates, of course, that some of the conditions which we have imposed on the operators appearing in the foregoing analysis are mutually incompatible. Now these conditions (or assumptions) are the following : 1) that the operators x_i, p_i be real, 2) that they be proper, and 3) that they satisfy the commutation relations (9.4). It actually turns out that the inconsistency (9.24) can be removed by relaxing any one of these three conditions only slightly. However, since these conditions seem so necessary from the point of view of our general dynamical theory, we shall relax any of them only with the greatest reluctance. Therefore, before doing anything drastic let us examine each one separately.

We certainly do not want to give up the commutation relations (9.4), since they form the backbone of the analogy between quantum mechanics and the canonical formalism of classical mechanics. Also, we certainly want the x_i and p_i to be proper operators, since only under such circumstances can we introduce the very convenient momentum or coordinate space representations. Therefore, all that remains to be questionned is the reality of the operators x_i and p_i. However, in a previous chapter we have already presented the arguments for the reality or Hermiticity of operators such as the x's and p's which correspond to classical quantities which are really

observable. This evidently leaves us in a dilemma.

We shall now show, however, that the dilemma can be completely resolved by regarding the momentum operators p_i not as real operators but as the limiting forms of complex operators whose imaginary or anti-Hermitian parts tend to zero. The coordinate operators x_i may be left alone in their originally assumed real forms.

We shall replace the differential form (9.12) of the momentum operator by

$$p = -i\hbar\nabla - \frac{1}{2}i\hbar\varepsilon\frac{r}{r^2} \tag{9.25}$$

where ε is a small positive quantity which is allowed to vanish after all algebraic work with the operators p and r is finished The eigenvalues of the operator (9.25) are still given by

$$p' = \hbar k' \tag{9.26}$$

but the wave function corresponding to the eigenvector $|p'\rangle$ is given by

$$\psi_{p'}(r) = \frac{1}{\sqrt{\Omega}} e^{ik'\cdot r - \frac{1}{2}\varepsilon r} \tag{9.27}$$

instead of by (9.14).

Since the operator (9.25) is not real, equation (9.23) does not follow from (9.9). We have, rather,

$$\langle p'|p^* = \langle p'|p' \tag{9.28}$$

where

$$p^* = -i\hbar\nabla + \frac{1}{2}i\hbar\varepsilon\frac{r}{r^2} = p + i\hbar\varepsilon\frac{r}{r^2} \tag{9.29}$$

The operator (9.25) still satisfies the commutation relations (9.4), but we may now write

$$
\begin{aligned}
i\hbar\,\delta_{ij} &= \langle p'|\, i\hbar\,\delta_{ij}\,|p'\rangle = \langle p'|[\alpha_i, p_j]|p'\rangle \\
&= \langle p'|\alpha_i p_j|p'\rangle - \langle p'|p_j \alpha_i|p'\rangle \\
&= p'_j \langle p'|\alpha_i|p'\rangle - \langle p'|\left(p_j^* - i\hbar \sum_\ell \frac{\alpha_\ell}{x_\ell}\right)\alpha_i|p'\rangle \\
&= (p_j - p'_j)\langle p'|\alpha_i|p'\rangle + i\hbar \langle p'|\sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell}|p'\rangle \\
&= i\hbar \langle p'|\sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell}|p'\rangle
\end{aligned}
\tag{9.30}
$$

The last quantity may be evaluated, with the aid of (9.27), as follows :

$$
\begin{aligned}
\langle p'|\sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell}|p'\rangle &= \sum_\nu \langle p'|\nu\rangle \sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell} \langle \nu|p'\rangle \\
&= \frac{\Delta\omega}{L^3} \sum_\nu \sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell} e^{-\varepsilon x'} = \frac{\varepsilon}{L^3} \int \sum_\ell \frac{\alpha_\ell \alpha_i}{x_\ell} e^{-\varepsilon x'} dx' \\
&= \frac{1}{3}\delta_{ij} \frac{\varepsilon^3}{L^3} \int x' e^{-\varepsilon x'} dx' = \frac{4\pi}{3}\delta_{ij} \frac{\varepsilon}{L^3} \int_0^\infty x'^3 e^{-\varepsilon x'} dx' \\
&= \frac{8\pi}{\varepsilon^3 L^3}\,\delta_{ij}\,.
\end{aligned}
\tag{9.31}
$$

Now, the normalisation condition (8.116) applied to the wave function (9.27) yields

$$
1 = \frac{1}{L^3}\int_\infty e^{-\varepsilon x} dx = \frac{4\pi}{L^3}\int x^2 e^{-\varepsilon x} dx = \frac{8\pi}{\varepsilon^3 L^3} \tag{9.32}
$$

from which we may infer

$$L = \frac{(8\pi)^{1/3}}{\epsilon} \tag{9.33}$$

and

$$\langle p'|\epsilon \frac{x_i x_j}{r}|p\rangle = \delta_{ij} \tag{9.34}$$

Equation (9.30) is now self-consistent, and we no longer have a contradiction.

There remains, however, a question about the orthogonality of the basic-vectors $|p'\rangle$; It will be recalled that the proof in chapter 7, of the orthogonality of a set of basic vectors defined by a complete set of operators, depended on the reality of those operators. Since the operator p is no longer real we need not expect the vectors $|p'\rangle$ to be now mutually orthogonal. Indeed, we have

$$\langle p''|p'\rangle = \sum_r \langle p''|r\rangle\langle r|p'\rangle = \int_\infty \psi^*_{p''}(r')\psi_{p'}(r')dr'$$

$$= \frac{1}{L^3}\int_\infty e^{i(k'-k'')\cdot r' - \epsilon r'} dr'$$

$$= \frac{\epsilon^3}{\pi}\int_{-1}^{1} d\alpha \int_0^\infty r'^2 dr' e^{i|k'-k''|r'\alpha - \epsilon r'}$$

$$= \frac{\epsilon^3}{4}\int_0^\infty r'^2 dr' \left[\frac{e^{i|k'-k''|r'\alpha}}{i|k'-k''|r'}\right]_{-1}^{1} e^{-\epsilon r'}$$

$$= \frac{\epsilon^3}{2|k'-k''|}\int_0^\infty r' e^{-\epsilon r'} \sin|k'-k''|r' dr' \tag{9.35}$$

But

$$\int_0^\infty x e^{-ax} \sin \beta x \, dx = -\frac{d}{da} \int_0^\infty e^{-ax} \sin \beta x \, dx$$

$$= -\frac{d}{da}\left[\frac{e^{-ax}(-a \sin \beta x - \beta \cos \beta x)}{a^2+\beta^2}\right]_0^\infty = -\frac{d}{da}\frac{\beta}{a^2+\beta^2}$$

$$= \frac{2a\beta}{(a^2+\beta^2)^2} \,. \tag{9.36}$$

Hence

$$\langle p''/p' \rangle = \frac{\epsilon^4}{(|k'-k''|^2+\epsilon^2)^2} \,. \tag{9.37}$$

Evidently

$$\lim_{\epsilon \to 0} \langle p''/p' \rangle = \begin{cases} 0 & \text{for } k' \neq k'' \\ \infty & \text{for } k' = k'' \end{cases} \tag{9.38}$$

The vectors $|p'\rangle$ therefore become orthogonal only in the limit $\epsilon \to 0$.

The reader has by this time doubtless inferred that the contradiction which was obtained in "equation" (9.24) is intimately related to the fact that the wave functions (9.14) are not, strictly speaking, integrable -more precisely, they are not quadratically integrable. The wave functions (9.27), on the other hand, are quadratically integrable, owing to the presence of the "damping factor" For this reason, the mathematics of quantum mechanics is sometimes presented in a framework in which only quadratically integrable wave functions are allowed. The mathematical theory is then spoken of as the theory of Hilbert space. However, in order to allow ourselves as much freedom as possible we do not here wish to restrict ourself within the confines of Hilbert space theory. The vector spaces with which we deal here may be regarded as generalizations of Hilbert space. We must, on the other hand, pay for the privilidge of the extra freedom allo-

wed by these generalizations, by being extra vigilant to detect contradictions like (9.24) so as not to be led into various mathematical traps.

The trick of turning a real operator into a complex operator with a small imaginary part will be frequently applied in the future when the operator has eigenvalues ranging over a continuum. It is easy to show that the wave functions corresponding to those eigenvectors of a proper real operator which come from an unbounded continuum must be necessarily non-quadratically-integrable. Let α be an operator which possesses such eigenvectors. Then

$$1 = \langle \rho' | \rho' \rangle = \sum_{\alpha',\gamma'} \langle \rho' | \alpha', \gamma' \rangle \langle \alpha', \gamma' | \rho' \rangle$$

$$= \Delta\omega \iint |\psi_{\alpha',\gamma'}(\rho')|^2 \delta(\alpha' - a) \delta(\gamma' - \gamma') \, d\alpha' \, d\gamma'$$

$$= \Delta\omega \, R \, \delta(\bar{\alpha}' - \bar{\alpha}) \delta(\bar{\gamma}' - \bar{\gamma}) |\psi_{\bar{\alpha}',\bar{\gamma}'}(\rho')|^2 \quad (9.39)$$

where R denotes the "volume" of the range of the eigenvalues α', γ', and the bars denote suitable average values. Now, R is an infinity which, in general, has the same number of dimensions as the infinitesimal $\Delta\omega$. Hence they cancel one another, leaving

$$|\psi_{\alpha',\gamma'}(\rho')|^2 \sim \frac{1}{\delta(\bar{\alpha}' - \bar{\alpha}') \, \delta(\bar{\gamma}' - \bar{\gamma}')} \quad (9.40)$$

This means that $|\psi_{\alpha',\gamma'}(\rho')|^2$ is actually infinitesimally small for all ρ' (except possible isolated points). But the normalization condition (8.116) shows that $|\psi_{\alpha',\gamma'}(\rho')|^2$ need be infinitesimally small only if the function $\psi_{\alpha',\gamma'}(\rho')$, apart from its infinitesimal constant factors, is non-quadratically integrable.

Returning now to the problem of the free particle, we note that since the Hamiltonian function is a function of the momenta only, the momentum eigenvectors $|p'\rangle$ are also energy eigenvectors. We have

$$H|p'\rangle = E'|p'\rangle \tag{9.41}$$

where $$E' = \frac{1}{2m} p'^2 \tag{9.42}$$

Consider the Schrödinger equation

$$i\hbar \frac{d}{dt}|t\rangle = H|t\rangle \tag{9.43}$$

Multiplying on the left by $\langle p'|$, we obtain*

$$i\hbar \frac{\partial}{\partial t}\langle p'|t\rangle = \langle p'|H|t\rangle = E'\langle p'|t\rangle. \tag{9.44}$$

The solution of (9.44) is

$$\langle p'|t\rangle = \langle p'|t_0\rangle\, e^{\frac{1}{i\hbar}E'(t-t_0)} \tag{9.45}$$

and we may write

$$|t\rangle = \sum_{p'} |p'\rangle\langle p'|t\rangle = \sum_{p'} |p'\rangle\langle p'|t_0\rangle e^{\frac{1}{i\hbar}E'(t-t_0)} \tag{9.46}$$

Equation (9.46) may be readily cast into wave function form by multiplying on the left by $\Delta\omega^{-1/2}\langle r'|$:

$$\psi(r';t) = \sum_{p'} \psi_{p'}(r')\langle p'|t_0\rangle e^{\frac{1}{i\hbar}E'(t-t_0)}$$

---

* Since we are not working with commutators here we may drop the imaginary part of the operator p and write immediately $\langle p'|H = \langle p'|\, E'$.

$$= \frac{1}{\sqrt{L^3}} \sum_{p'} \langle p'/t_0 \rangle e^{-\frac{i}{\hbar} E' t_0} e^{i/\hbar (p' \cdot r - E' t)}$$

$$= \int f(p') e^{i/\hbar (p' \cdot r - E' t)} d_3 p' \qquad (9.47)$$

where

$$f(p') = \frac{\delta(p' - p')}{\sqrt{L^3}} \langle p'/t_0 \rangle e^{\frac{i}{\hbar} E' t_0} \qquad (9.48)$$

Equation (9.47) is readily seen to be the general solution of the wave function form of the Shrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \psi(r,t) = H \psi(r,t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(r,t) \qquad (9.49)$$

The function $f(p')$ is completely arbitrary.

Equation (9.47) represents the general wave function as a superposition of plane waves, each having an angular frequency:

$$\omega = \frac{E'}{\hbar} \qquad (9.50)$$

and an angular wavelength

$$\lambda = \hbar / p' \,. \qquad (9.51)$$

The phase velocity of these waves is evidently

$$v_p = \lambda \omega = \frac{E'}{p'} = \frac{1}{2m} p' \,. \qquad (9.52)$$

Suppose now that the function $f(p')$ has the form

$$f(p') = F(p' - p_0) e^{-\frac{i}{\hbar} (p' \cdot r_0 - E' t_0)} \qquad (9.53)$$

Then

$$\psi(r,t) = \int F(p'-p_0)\, e^{i/\hbar\,[p'\cdot(r-r_0) - E'(t-t_0)]}\, d_3p' \qquad (9.54)$$

If the function F is a smoothly varying function which is peaked at the origin and which tapers off to zero on either side so that F is non-vanishing essentially only in a region having roughly a width Δp, say, centered at the origin, then we may show that the region of space in which ψ is non-vanishing is also more or less bounded. Let us consider the time $t = t_c$. Then at the point $r' = r_0$, the exponential has the value unity and all the infinitesimal contributions to the integral (9.54) add in phase. $\psi(r', t_c)$ is therefore peaked at r_0. The infinitesimal contributions to the integral at a point distant δr from r_0 can add up out of phase and hence cancel one another only if the integrand has significant magnitudes for values of p' satisfying

$$|(p'-p_0)\cdot\delta r| \geq \pi\hbar \qquad (9.55)$$

$\psi(r', t_c)$ remains large therefore within a region of dimensions Δx satisfying

$$\Delta p\, \Delta x \simeq h \simeq \hbar \qquad (9.56)$$

Equation (9.56) is seen to be a less rigorous form of the Heisenberg uncertainty relation (8.147).

The wave function above is said to describe a "wave packet" centered at the point r_0 at the time t_0. The width of the wave packet at time t_0 is Δx. At later times; as we shall presently see, the wave packet will spread out over dimensions larger than Δx. It will also move. Before examining the rate of spreading, let us find the velocity of motion of the packet. The peak of the packet at any instant is located at the point at which the phase of the exponential in (9.54) has an extremum with respect to p' at $p' = p_0$. For at

this point, the main infinitesimal contributions to the integral add up constructively. This point is evidently given by

$$0 = \left[\frac{\partial}{\partial p'}\left\{p'(r-r_0) - E'(t-t_0)\right\}\right]_{p'=p_0}$$

$$= (r'-r_0) - \frac{1}{m}p_0(t-t_0) \qquad (9.57)$$

The velocity of the wave packet is therefore

$$v_g = \frac{r'-r_0}{t-t_0} = \frac{1}{m}p_0 \qquad (9.58)$$

The velocity v_g is known as the group velocity of the probability wavesn and is seen to be exactly equal to the velocity of the particle when it has the momentum p_0.

The wave function (9.54) represents a quantum state in which the expectation value of the momentum is p_0, but in which there is a certain amount of uncertainty in the precision with which the momentum is known, reflected by the width Δp of the function F. This wave function therefore describes a situation which is always encountered in reality, for no observables can be measured with absolute precision. The behavior of the wave function has much in common with the behavior (in classical terms) of the particle which it describes. The description of material particles by means of probability waves $e^{\frac{i}{\hbar}p\cdot r}$ and wave packets was first introduced into physics by L. de Broglie.

In order to calculate the rate of spreading of a given wave packet, we may observe that the width of the packet and the width of the function F may be defined rigorously as the root mean square deviations Δx_i and Δp_i respectively of the coordinates and momenta of the particle. Let us therefore obtain the time dependence of Δx_i. In the equations which follow we shall not use the implicit summation convention with respect to

the indices i. The mean value of the momentum, denoted in (9.54) ff. by p_0, has components given by

$$p_{oi} = \langle t | p_i | t \rangle \tag{9.59}$$

and is independent of time :

$$\dot{p}_{oi} = \frac{1}{i\hbar} \langle t | [p_i, H] | t \rangle = 0. \tag{9.60}$$

The mean value of the particle position, on the other hand, depends linearly on the time

$$\overline{x_i} = \langle t | x_i | t \rangle, \tag{9.61}$$

$$\dot{\overline{x}}_i = \frac{1}{i\hbar} \langle t | [x_i, H] | t \rangle = \frac{1}{m} \langle t | p_i | t \rangle = \frac{1}{m} p_{oi}, \tag{9.62}$$

$$\ddot{\overline{x}}_i = \frac{1}{m} \dot{p}_{oi} = 0, \tag{9.63}$$

so that

$$\overline{x_i} = x_{oi} + \frac{1}{m} p_{oi} t. \tag{9.64}$$

The $\overline{x_i}$ may be identified with the coordinates of the peak of the wave packet. Hence writing

$$\overline{x}_i = x_{oi} + v_g t, \tag{9.65}$$

we are again led to expression (9.53) for the group velocity.

The mean square of a momentum component is

$$\overline{p_i^2} = \langle t | p_i^2 | t \rangle. \tag{9.66}$$

Like p_{oi}, it is time independent. The mean square of a coordinate,

however, depends quadratically on the time :

$$\overline{x_i^2} = \langle t | x_i^2 | t \rangle, \tag{9.67}$$

$$\overline{\dot{x}_i} = \frac{1}{i\hbar} \langle t | [x_i^2, H] | t \rangle = \frac{1}{m i\hbar} \langle t | \{x_i, p_i\} | t \rangle$$

$$= \frac{1}{m} \overline{\{x_i, p_i\}}, \tag{9.68}$$

$$\overline{\ddot{x}_i} = \frac{1}{i\hbar m} \langle t | \{x_i, p_i\}, H | t \rangle = \frac{2}{m^2} \langle t | p_i^2 | t \rangle$$

$$= \frac{2}{m^2} \overline{p_i^2}, \tag{9.69}$$

$$\overline{\dddot{x}_i} = \frac{2}{m^2} \overline{p_i^2} = 0, \tag{9.70}$$

so that

$$\overline{x_i^2} = \overline{(x_i^2)_0} + \frac{1}{m} \overline{\{x_i, p_i\}_0} t + \frac{1}{m} \overline{p_i^2} t^2. \tag{9.71}$$

Combining eqns. (9.64) and (9.71), we now have

$$\Delta x_i^2 = \overline{x_i^2} - \overline{x_i}^2$$

$$= \overline{(x_i^2)_0} - x_{0i}^2 + \frac{1}{m} \left[\overline{\{x_i, p_i\}_0} - 2 x_{0i} p_{0i} \right] t$$

$$+ \frac{1}{m^2} \left(\overline{p_i^2} - p_{0i}^2 \right) t^2$$

$$= \overline{(\Delta x_i^2)_0} + \frac{1}{m} \left[\overline{\{x_i, p_i\}_0} - 2 x_{0i} p_{0i} \right] + \frac{1}{m^2} \Delta p_i^2 t^2. \tag{9.72}$$

Let us now seek for a wave function for which, at time $t = t_0$, the equality sign holds in the Heisenberg uncertainty relation (8.147). Referring back to the fundamental equations (7.148) and (7.149) from which (8.147) was derived, we see that the equality sign can hold only if the vectors $(x_i - x_{0i}) \mid t_0\rangle$ and $(p_i - p_{0i}) \mid t_0\rangle$ are parallel. In wave function language this condition becomes

$$(r - r_0)\,\psi(r, t_0) = a\,(p - p_0)\,\psi(r, t_0) \qquad (9.73)$$

or

$$-i\hbar\,\nabla\psi(r, t_0) = \frac{1}{a}\,(r - r_0 + a p_0)\,\psi(r, t_0), \qquad (9.74)$$

where a is an arbitrary constant and where we have now added the requirement that the wave packet be spherically symmetric. The general solution of (9.74) is

$$\psi(r, t_0) = A\,e^{\frac{i}{2\hbar a}(r - r_0 + a p_0)^2}, \qquad (9.75)$$

where the coefficient A is to be chosen so as to normalize the wave function.

Referring once again to eq. (7.149) we see that in order for the equality sign to hold in (8.147), we must also have

$$\langle t_0 \mid \{x_i, p_i\} - 2 x_{0i} p_{0i} \mid t_0\rangle = \overline{\{x_i, p_i\}}_0 - 2 x_{0i} p_{0i} \overset{(9.76)}{=} 0$$

But, using (9.75), we have

$$\overline{\{x_i, p_i\}}_0 = -i\hbar \int_{-\infty}^{\infty} \psi^*(r, t_0)\left(x_i \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_i} x_i\right)\psi(r, t_0)\, d^3 r$$

$$= -i\hbar C \int_{-\infty}^{\infty} e^{-\frac{i}{2\hbar a}(x_i - x_{0i} + a^* p_{0i})^2} \left(x_i \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_i} x_i\right) e^{\frac{i}{2\hbar a}(x_i - x_{0i} + a p_{0i})^2} dx_i \qquad (9.77)$$

where C is a certain constant introduced as a result of integration

over two of the coordinates. Now a is so far an arbitrary complex number. In order that the wave packet be confined to a limited region of space it is evident that we must have

$$\varphi_m(a) < 0 \tag{9.78}$$

As a result of this condition, we may integrate by parts in (9.77), obtaining

$$\overline{\{x_0 p_0\}} = C \int_{-\infty}^{\infty} e^{-\frac{1}{2\hbar a^*}(x_0 - x_{0c} + a^* p_{0c})^2} \left[\frac{1}{a^*} x_0 (x_0 - x_{0c} - a^* p_{0c}) \right.$$

$$\left. + \frac{1}{a^*}(x_0 - x_{0c} + a^* p_{0c}) x_0 \right] e^{\frac{1}{2\hbar a}(x_0 - x_{0c} - a p_{0c})^2} dx_0$$

$$= \left(\frac{1}{a} + \frac{1}{a^*}\right)(\Delta x_0^2)_0 + 2 x_{0c} p_{0c} \tag{9.79}$$

In order for (9.76) to hold we see that we must have

$$\frac{1}{a} + \frac{1}{a^*} = 0, \tag{9.80}$$

which means that a must be pure imaginary. Remembering condition (9.78), we write

$$a = -i\alpha \qquad \alpha > 0 \tag{9.81}$$

and

$$\psi(r, t_0) = A e^{-\frac{1}{2\hbar\alpha}(r - r_0 - i\alpha p_0)^2}. \tag{9.82}$$

The wave function is thus seen to follow a Gaussian curve. Since equation (9.76) and the equality sign in (8.147) now hold, equation

(9.72) reduces to

$$\Delta x_i^2 - (\Delta x_i^2)_0 = \frac{\hbar^2}{4m^2(\Delta x_i^2)_0} t^2 \tag{9.83}$$

The coefficient A in (9.82) may be partially determined from the normalization requirement (8.116) with the aid of the following definite integral (cf. Appendix (2.52)) :

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} . \tag{9.84}$$

We have

$$\begin{aligned} 1 &= \int_\infty \psi^*(r, t_0)\, \psi(r, t_0)\, d_3 r \\ &= |A|^2 e^{\alpha/\hbar p_0^2} \int_\infty e^{-\frac{1}{\hbar\alpha}(r - r_0)^2} d_3 r \\ &= |A|^2 e^{\alpha/\hbar p_0^2} (\pi \hbar \alpha)^{3/2}, \end{aligned} \tag{9.85}$$

from which we infer

$$A = (\pi \hbar \alpha)^{-3/4} e^{-\alpha/2\hbar p_0^2} . \tag{9.86}$$

Differentiating equation (9.84) with respect to a, we obtain

$$\int_{-\infty}^{\infty} x^2 e^{-\alpha x^2} dx = -\frac{d}{d\alpha} \sqrt{\frac{\pi}{\alpha}} = \frac{1}{2} \sqrt{\frac{\pi}{\alpha^3}} . \tag{9.87}$$

Using this equation we may establish a relation between the constant and the mean square deviation $(\Delta x_i^2)_0$:

$$(\Delta x_i^2)_0 = \int_\infty \psi^*(r, t_0)\, (x_i - x_{0i})^2\, \psi(r, t_0)\, d^3 r$$

$$= |A|^2 e^{\alpha/\hbar p_0^2} \pi\hbar\alpha \int_{-\infty}^{\infty} (x_i - x_{0i})^2 e^{-\frac{1}{\hbar\alpha}(x_i - x_{0i})^2} dx_i$$

$$= \frac{\hbar\alpha}{2} |A|^2 e^{\alpha/\hbar p_0^2} (\pi\hbar\alpha)^{3/2} = \frac{\hbar\alpha}{2}. \qquad (9.88)$$

We may now write (9.82) in the form

$$\psi(r, t_0) = [2\pi(\Delta x^2)_0]^{-3/4} e^{-\frac{1}{4(\Delta x^2)_0}(r - r_0)^2 + \frac{i}{\hbar} p_0 \cdot (r - r_0)}, \qquad (9.89)$$

where we have dropped the subscript i from $(\Delta x_i^2)_0$, owing to symmetry.

Because of the symmetry between coordinates and momenta at the time $t = t_0$, we may expect that the momentum distribution in the wave packet has the same form as the coordinate distribution, i.e. that the function $f(p')$ of (9.48) has the form*:

$$f(p') e^{-\frac{i}{\hbar} E' t_0} = B e^{-\frac{1}{2\beta}(p' - p_0 + \frac{E'}{v_0})^2}, \qquad (9.90)$$

where B and β are suitable constants, analogous to the A and α of (9.82). In fact, on the basis of (9.88) we may even infer

$$\beta = \frac{2}{\hbar^2} \Delta p_i^2 = \frac{\hbar^2}{2(\Delta x_i^2)_0}. \qquad (9.91)$$

Equations (9.90; 91) may be verified by direct substitution in (9.48). Using the Fourier transform relation

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2a}x^2 + ikx} dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2a}(x - ak)^2 - \frac{a}{2}k^2} dx$$

---

* The + sign must be taken in front of the term $\frac{i}{\hbar} E' t_0$ in (9.90) instead of the - sign since the differential form of p is $\frac{\hbar}{i}\frac{\partial}{\partial p}$ while that of p is $-i\hbar\frac{\partial}{\partial x}$.

$$= \int_{-\infty - \frac{i}{2}\alpha k}^{\infty - \frac{i}{2}\alpha k} e^{-\frac{1}{2\alpha}z^2 - \frac{\alpha}{2}k^2} dz = e^{-\frac{\alpha}{2}k^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2\alpha}z^2} dz$$

$$= \sqrt{2\pi\alpha}\, e^{-\frac{\alpha}{2}k^2}, \qquad (9.92)$$

we have

$$\psi(r,t_0) = B e^{\frac{i}{2\hbar\alpha}r_0^2} \int_{-\infty} e^{-\frac{1}{2\hbar\beta}(p'-p_0)^2 - \frac{i}{\hbar}(p'-p_0)r_0} e^{\frac{i}{\hbar}p'r} d_3p'$$

$$= B e^{\frac{i}{2\hbar\alpha}r_0^2 + \frac{i}{\hbar}p_0 r} \int_{-\infty} e^{-\frac{1}{2\hbar\beta}(p'-p_0)^2 - \frac{i}{\hbar}(p'-p_0)\cdot r - r_0} d_3p'$$

$$= (2\pi\hbar\beta)^{3/2} B e^{\frac{i}{2\hbar\alpha}r_0^2 + \frac{i}{\hbar}p_0 r_0 - \frac{i}{2\hbar\alpha}(r-r_0)^2}$$

(9.93)

Comparison with (9.89) yields, for the constant B,

$$B = \left[2\pi(\Delta x^2)_0\right]^{-3/4} (2\pi\hbar\beta)^{-3/2} e^{\frac{i}{2\hbar\alpha}r_0^2 - \frac{i}{\hbar}\delta} \qquad (9.94)$$

We may now obtain the expression for the wave function at an arbitrary time t:

$$\psi(r,t) = \left[2\pi(\Delta x^2)_0\right]^{-3/4} (2\pi\hbar\beta)^{-3/2} \int_{-\infty} e^{\frac{i}{2\hbar\alpha}(p'-p_0)^2 + \frac{i}{\hbar}p'(r-r_0) - \frac{i}{\hbar}E'(t-t_0)} d_3p'$$

$$= \left[2\pi(\Delta x^2)_0\right]^{-3/4} (2\pi\hbar\beta)^{-3/2} \int_{-\infty} e^{-\frac{i}{2\hbar\alpha}\left[\left(1 - \frac{i\beta}{m}(t-t_0)\right)p'^2 - 2\left(r_0 + \frac{i\beta}{m}(r-r_0)\right)p' + p_0^2\right]} d_3p'$$

$$= \left[2\pi(\Delta x^2)_0\right]^{-3/4} (2\pi\hbar\beta)^{-3/2} \int_{-\infty} e^{-\frac{1}{2\hbar\alpha}\left[\left(1 + \frac{i\beta}{m}(t-t_0)\right)p'^2 - 2\left(r_0 + \frac{i\beta}{m}(r-r_0)\right)p' + p_0^2\right]} d_3p'$$

$$= \left\{2\pi(\Delta x^2)_0\left[1 + \frac{i\beta}{m}(t-t_0)\right]^2\right\}^{-3/4} e^{-\frac{i}{2\hbar\alpha}\left(r_0 - \frac{p_0}{m}(t-t_0)\right)^2 \frac{1}{1+\frac{i\beta}{m}(t-t_0)} + \frac{i}{\hbar}p_0(r-r_0) - \frac{i}{\hbar}E(t-t_0)}$$

(9.95)

The exponent in (9.95) may be rearranged thus

$$-\frac{1}{2\hbar\beta}\,\frac{-2i\beta\left[p_0\cdot(r-r_0)-\frac{1}{2m}p_0^2(t-t_0)\right]\left[1+i\frac{\beta}{m}(t-t_0)\right]-2\frac{\beta^2}{m}p_0\cdot(r-r_0)(t-t_0)+\frac{\beta^2}{m^2}p_0^2(t-t_0)^2+\beta^2(r-r_0)^2}{1+i\frac{\beta}{m}(t-t_0)}$$

$$=\frac{i}{\hbar}\left[p_0\cdot(r-r_0)-E_0(t-t_0)\right]-\frac{\beta}{2\hbar}\,\frac{\left[r-r_0-\frac{1}{m}p_0(t-t_0)\right]^2}{1+i\frac{\beta}{m}(t-t_0)} \qquad (9.96)$$

Hence

$$\psi(r,t)=\left\{2\pi(\Delta x^2)_0\left[1+i\frac{\hbar(t-t_0)}{2m(\Delta x^2)_0}\right]^2\right\}^{-3/4} e^{\frac{i}{\hbar}\left[p_0\cdot(r-r_0)-E(t-t_0)\right]}$$

$$\times e^{-\frac{1}{4(\Delta x^2)_0}\,\frac{\left[r-r_0-\frac{p_0}{m}(t-t_0)\right]^2}{1+i\frac{\hbar(t-t_0)}{2m(\Delta x^2)_0}}} \qquad (9.97)$$

(9.97) is seen to reduce to (9.89) when $t = t_0$.

Problem XLVII : Let $\xi = \frac{1}{2}\{x,p\}$. Show that the eigenvalues ξ' range $-\infty$ to $+\infty$ and that the complete set of orthonormal eigenvectors $|\xi'\rangle$ has the representation

$$\langle x'|\xi'\rangle = \left[\frac{\Delta\omega}{4\pi\hbar\,\delta(\xi''-\xi')}\right]^{1/2}\frac{e^{\frac{i}{\hbar}\xi' l_{op}\frac{|x'|}{\lambda}}}{\sqrt{|x'|}},$$

where λ is an arbitrary positive constant.

b) The harmonic oscillator.

The Hamiltonian function is (see (1.116))

$$H=\frac{1}{2m}\left(p^2+m^2\omega^2p^2\right). \qquad (9.98)$$

The canonical equations of motion are

$$\dot{p}=\frac{1}{i\hbar}[p,H]=\frac{1}{m}p, \qquad (9.99)$$

$$\dot{p} = \frac{1}{i\hbar}[p, H] = -m\omega^2 \rho, \qquad (9.100)$$

Leading to

$$\ddot{\rho} + \omega^2 \rho = 0 \ , \quad \ddot{p} + \omega^2 p = 0 \qquad (9.101)$$

The solutions of these equations are evidently

$$\rho(t) = \rho(t') \cos\omega(t-t') + \frac{1}{m\omega} p(t') \sin\omega(t-t') \qquad (9.102)$$

$$p(t) = -m\omega \rho(t') \sin\omega(t-t') + p(t') \cos\omega(t-t') \qquad (9.103)$$

from which we may infer

$$[\rho(t), \rho(t')] = -\frac{i\hbar}{m\omega} \sin\omega(t-t') \qquad (9.104)$$

It may be observed that the quantity on the right of (9.104) is equal to $-\frac{i\hbar}{m}$ times the difference of the retarded and advanced Green's functions for the harmonic oscillator (see (2.85)). In the limit $\omega \to 0$ (9.104) reduces to the 1-dimensional form of (9.7). The uncertainty relation for the root mean square deviations of successive measurements of ρ takes, in this case, the form

$$\Delta\rho(t)\, \Delta\rho(t') \geq \frac{\hbar}{2m\omega} |\sin\omega(t-t')| \qquad (9.105)$$

If the two measurements are made at a time interval of precisely an integral number of periods, then they do not interfere with each other. This phenomenon is peculiar to the harmonic oscillator and reflects the fact that no matter what initial momentum the particle has it will always return to its starting position after one period. This situation is also responsible for the fact that wave packets for a harmonic oscillator, unlike those for the free particle, do not spread out indefinitely in time. On the contrary, they pulsate, regaining their original

[illegible] we obtain the transformation matrix

form after each period. We shall not, however, go into the detailed analysis of harmonic oscillator wave packets.

Let us introduce the following operators :

$$a = \frac{1}{\sqrt{2m\hbar\omega}}(p + im\omega p), \quad a^* = \frac{1}{\sqrt{2m\hbar\omega}}(p - im\omega p) \tag{9.106}$$

These operators satisfy the equations

$$\dot{a} = i\omega a \quad , \quad \dot{a}^* = -i\omega a^* , \tag{9.107}$$

of which the general solutions are

$$a = a_S e^{i\omega(t-t_0)}, \quad a^* = a_S^* e^{-i\omega(t-t_0)} \tag{9.108}$$

a_S and a_S^* being the Schrödinger form of the operators. We shall now work in the Schrödinger representation and drop the subscripts S for brevity.

Let us first observe the following relations :

$$\hbar\omega\, a\, a^* = \frac{1}{2m}(p + im\omega p)(p - im\omega p)$$

$$= \frac{1}{2m}(p^2 + m^2\omega^2 p^2) + \frac{1}{2}\omega[p, p]$$

$$= H - \frac{1}{2}\hbar\omega \tag{9.109}$$

and similarly

$$\hbar\omega\, a^* a = H + \frac{1}{2}\hbar\omega, \tag{9.110}$$

and hence

$$[a^*, a] = -1 \tag{9.111}$$

We may use these relations to determine the eigenvectors and eigenvalues of the Hamiltonian operator. An eigenvector of the operator aa^* is evidently also an eigenvector of H , and the eigenvalues of H are determined by the eigenvalues of aa^*. Hence we may work with the operators a and a^* alone.

Let us introduce the definition

$$n \equiv a a^* \tag{9.112}$$

Then, using postulate 17) of chapter 7), we have the following condition on the eigenvalues of n :

$$n' = \langle n'/n/n' \rangle = (\langle n'/a)(a^*/n')) \geq 0 \tag{9.113}$$

The equality sign can hold in (9.113) only if $a^*/n') = 0$. If $a^*/n') \neq 0$, then, using the first of the commutation relations

$$\left.\begin{aligned} [n, a^*] &= [aa^*, a^*] = [a, a^*]a^* = -a^* \\ [n, a] &= [aa^*, a] = a[a^*, a] = a \end{aligned}\right\} \tag{9.114}$$

we have

$$n a^*/n') = ([n, a^*] + a^* n)/n') = (n' - 1) a^*/n') \tag{9.115}$$

which says that $a^*/n')$ is an eigenvector of n corresponding to the eigenvalue $n' - 1$. Similarly, if $a^{*2}/n') \neq 0$ it is an eigenvector corresponding to the eigenvalue $n' - 2$. Continuing in this manner, repeatedly applying the operator a^*, we see that n must have a series of eigenvalues n', $n' - 1$, $n' - 2$, ... Unless the equality sign in (9.113) holds at some point in this procedure, the eigenvalues of n will eventually become negative, in contradiction

with (9.113). The lowest eigenvalue of n is therefore that for which the equality sign in (9.113) holds, namely zero. We have

$$a^* |0\rangle = 0. \tag{9.117}$$

Using the second of the commutation relations (9.114), we have

$$n a |n'\rangle = \left([n, a] + an\right)|n'\rangle = (n'-1) a |n'\rangle \tag{9.118}$$

which says that $a|n'\rangle$ is an eigenvector of n corresponding to the eigenvalue $n'-1$. We see, therefore, that the eigenvalues of n are the series of integers,

$$n' = 0, 1, 2, \ldots, \tag{9.119}$$

the corresponding eigenvectors being obtained by repeated application of a on $|0\rangle$. Since

$$\langle 0|a^{*n'} a^{n'}|0\rangle = \langle 0|a^{*n'-1}\left[a^*, a^{n'}\right]|0\rangle + \langle 0|a^{*n'-1} a^* a^{n'}|0\rangle$$

$$= n' \langle 0|a^{*n'-1} a^{n'-1}|0\rangle = \ldots = n'(n'-1)\ldots 2 \cdot 1 \langle 0|0\rangle = n'!, \tag{9.120}$$

we have the normalization condition

$$|n'\rangle = \frac{1}{\sqrt{n'!}}\, a^{n'}|0\rangle. \tag{9.121}$$

We may now write

$$a|n'-1\rangle = \frac{1}{\sqrt{(n'-1)!}}\, a^{n'}|0\rangle = \sqrt{n'}\,|n'\rangle \tag{9.122}$$

and

$$a^*|n'\rangle = \frac{1}{\sqrt{n'}} a^* a |n'-1\rangle = \frac{1}{\sqrt{n'}} \left([a^*, a] + a a^*\right)|n'-1\rangle$$

$$= \frac{1}{\sqrt{n'}} (n'+1)|n'-1\rangle = \sqrt{n'}\,|n'-1\rangle \qquad (9.123)$$

The matrix elements of the operators a and a^* are therefore given by

$$\langle n'|a|n'-1\rangle = \sqrt{n'}, \quad \langle n'-1|a^*|n'\rangle = \sqrt{n'} \qquad (9.124)$$

all other matrix elements vanishing.

Since the Hamiltonian operator can be written in the form (see (9.109))

$$H = \hbar\omega\left(n + \tfrac{1}{2}\right) \qquad (9.125)$$

we see that the energy eigenvalues of the harmonic oscillator are

$$E_{n'} = \hbar\omega\left(n' + \frac{1}{2}\right), \qquad (9.126)$$

in complete agreement with the result (5.12) of the old quantum theory. The matrix elements of p and p can be readily obtained if we remember that

$$p = -i\sqrt{\frac{\hbar}{2m\omega}}\left(a - a^*\right) \qquad (9.127)$$

$$p = \sqrt{\frac{m\hbar\omega}{2}}\left(a + a^*\right) \qquad (9.128)$$

Using (9.124), we obtain

$$\langle n'-1|p|n'\rangle = -\langle n'|p|n'-1\rangle = i\sqrt{\frac{\hbar}{2m\omega}}\sqrt{n'}, \qquad (9.129)$$

$$\langle n'-1|p|n'\rangle = \langle n'|p|n'-1\rangle = \sqrt{\frac{m\omega\hbar}{2}}\sqrt{n'} \qquad (9.130)$$

all other matrix elements vanishing. If the harmonic oscillator carries a charge e, its electric dipole moment is $\mu = ex$, of which the matrix elements are

$$\langle n'-1|\mu|n'\rangle = -\langle n'|ex|n'-1\rangle = -e\sqrt{\frac{h}{2\omega}}\sqrt{n'} \qquad (9.131)$$

Equation (9.131) agrees, up to an arbitrary phase factor $\mp i$, with the result (5.48) inferred from the old quantum theory on the basis of the Correspondence Principle.

Since the vectors $|n'\rangle$ are eigenvectors of H as well as of n', the energy eigenvalues E_n, may be used as labels just as well as the intergers n'. If we now transform back to the Heisenberg representation, we may write, remembering equations (9.108),

$$\langle E_{n'}|x|E_{n'-1}\rangle = \sqrt{\frac{h}{2\omega}}\sqrt{n'}\, e^{i\omega(t-t')} = \sqrt{\frac{h}{2\omega}}\sqrt{n'}\, e^{\frac{i}{h}(E_{n'}-E_{n'-1})(t-t')} \qquad (9.132)$$

$$\langle E_{n'-1}|x^*|E_{n'}\rangle = \sqrt{\frac{h}{2\omega}}\sqrt{n'}\, e^{-i\omega(t-t')} = \sqrt{\frac{h}{2\omega}}\sqrt{n'}\, e^{-\frac{i}{h}(E_{n'}-E_{n'-1})(t-t')} \qquad (9.133)$$

$$\langle E_{n'-1}|p|E_{n'}\rangle = -\langle E_{n'}|p|E_{n'-1}\rangle^* = -i\sqrt{\frac{h\omega}{2}}\sqrt{n'}\, e^{i\omega(t-t')} \qquad (9.134)$$

$$\langle E_{n'-1}|p|E_{n'}\rangle = -\langle E_{n'}|p|E_{n'-1}\rangle^* = -i\sqrt{\frac{h\omega}{2}}\sqrt{n'}\, e^{-i\omega(t-t')} \qquad (9.135)$$

These equations have, of course, the same form as (6.6).

We see from equations (9.122, 123) that if any function of the operators a, a^* is applied to any linear combination of the vectors $|n'\rangle$, the result will again be a linear combination of the vectors $|n'\rangle$. Since x and p are expressible in terms of a and a^* (see (9.127, 128)), this is also true of any function of x and p. Hence, if we work in the vector space generated by the vectors $|n'\rangle$ we shall never have to enlarge that vector space as a result of any of our dynamical equations. In this sense, then, the vectors $|n'\rangle$ form a complete set, and the operator n, or H, by itself forms a complete set of commuting proper real operators. Now, it is

often convenient to work in the configuration space representation in which ρ is diagonal. The completeness of the $|n'\rangle$'s implies that, unless there is a hidden incompatibility between the formalism of this section and the configuration space formalism of chapter 8, the wave functions corresponding to the $|n'\rangle$'s will form a complete orthonormal set of functions in ρ'-space. We shall now verify this implication by direct construction.

Equation (9.117), written in terms of ρ and p, has the form

$$(p - im\omega\rho)|0\rangle = 0 \tag{9.136}$$

Multiplying on the left by $\langle\rho'|$, we obtain

$$-i\hbar\left(\frac{\partial}{\partial\rho'} + \frac{m\omega}{\hbar}\rho'\right)\langle\rho'|0\rangle = 0 \tag{9.137}$$

of which the general solution is

$$\langle\rho'|0\rangle = A e^{-\frac{m\omega}{2\hbar}\rho'^2} \tag{9.138}$$

The constant A is fixed by the normalization condition. We have

$$1 = \sum_{\rho'}\langle 0|\rho'\rangle\langle\rho'|0\rangle = \Delta\omega^{-1}\int_{-\infty}^{\infty}|\langle\rho'|0\rangle|^2 d\rho'$$
$$= \Delta\omega^{-1}|A|^2\int_{-\infty}^{\infty} e^{-\frac{m\omega}{\hbar}\rho'^2} d\rho' = \Delta\omega^{-1}|A|^2\sqrt{\frac{\pi\hbar}{m\omega}} \tag{9.139}$$

We may evidently choose

$$A = \Delta\omega^{1/2}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \tag{9.140}$$

so that

$$\langle\rho'|0\rangle = \Delta\omega^{1/2}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}\rho'^2}. \tag{9.141}$$

Let us introduce

$$y = \sqrt{\frac{m\omega}{\hbar}}\,\rho' \tag{9.142}$$

Then, in differential form the operator a becomes

$$a = \frac{1}{\sqrt{2m\hbar\omega}}\left(-i\hbar\sqrt{\frac{m\omega}{\hbar}}\frac{\partial}{\partial y} + im\omega\sqrt{\frac{\hbar}{m\omega}}\,y\right)$$

$$= -\frac{i}{\sqrt{2}}\left(\frac{\partial}{\partial y} - y\right), \tag{9.143}$$

and we may write

$$\langle \rho'/0\rangle = A e^{-\frac{1}{2}y^2}, \tag{9.144}$$

and

$$\langle \rho'/n'\rangle = \frac{1}{\sqrt{n'!}}\langle \rho'/a^{n'}/0\rangle$$

$$= A\,\frac{(-i)^{n'}}{\sqrt{2^{n'}n'!}}\left(\frac{\partial}{\partial y} - y\right)^{n'} e^{-\frac{1}{2}y^2}. \tag{9.145}$$

Now observe that

$$e^{\frac{1}{2}y^2}\frac{\partial}{\partial y}e^{-\frac{1}{2}y^2} = \frac{\partial}{\partial y} - y \tag{9.146}$$

Hence we may write

$$\langle \rho'/n'\rangle = A\,\frac{(-i)^{n'}}{\sqrt{2^{n'}n'!}}\,e^{\frac{1}{2}y^2}\frac{\partial^{n'}}{\partial y^{n'}}e^{-y^2}$$

$$= A\,\frac{i^{n'}}{\sqrt{2^{n'}n'!}}\,e^{-\frac{1}{2}y^2}\,h_{n'}(y), \tag{9.147}$$

where

$$h_{n'}(y) = (-1)^{n'} e^{y^2}\frac{\partial^{n'}}{\partial y^{n'}}e^{-y^2} \tag{9.148}$$

The functions $h_{n'}(y)$ are known as Hermite polynomials :

$$\left.\begin{aligned} h_0(y) &= 1 \\ h_1(y) &= 2y \\ h_2(y) &= 4y^2 - 2 \\ h_3(y) &= 8y^3 - 12y \\ h_4(y) &= 16y^4 - 48y^2 + 12 \end{aligned}\right\} \qquad (9.149)$$

. . . . . .

Consider the function

$$f(y,t) = \sum_{n=0}^{\infty} \frac{1}{n!} h_n(y)\, t^n \qquad (9.150)$$

Inserting the definition (9.148) into this equation and using equation ( ) of the Appendix, we may write $f(y,t)$ in the form

$$f(y,t) = e^{y^2} \sum \frac{(-1)^n}{n!} t^n \frac{\partial^n}{\partial y^n} e^{-y^2} = e^{y^2} e^{-t\frac{\partial}{\partial y}} e^{-y^2}$$

$$= e^{y^2} e^{-(y-t)^2} = e^{-t^2+2yt} \qquad (9.151)$$

e^{-t^2+2yt} is called the generating function for the Hermite polynomials.

We are now in a position to verify explicitly the completeness of the set of basic vectors $|n'\rangle$. We have

$$\sum_{n'} \langle \rho'' | n' \rangle \langle n' | \rho' \rangle = |A|^2 e^{-\frac{1}{2}(y'^2+y''^2)} \sum_{n'} \frac{1}{2^{n'} n'!} h_{n'}(y) h_{n'}(y'')$$

$$= \Delta\omega \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-\frac{1}{2}(y'^2 - y''^2)} \sum_{n'} \frac{(-1)^{n'}}{2^{n'} n'!} h_{n'}(y') \left(\frac{\partial}{\partial y'}\right)^{n'} e^{-y''^2}$$

$$= \Delta\omega \left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-\frac{1}{2}(y'^2 - y''^2)} e^{-\frac{1}{4}\frac{\partial^2}{\partial y'^2} - y'\frac{\partial}{\partial y''}} e^{-y''^2} \qquad (9.152)$$

But

$$e^{-\frac{1}{4}\frac{\partial^2}{\partial y''^2} - y'\frac{\partial}{\partial y''}} e^{-y''^2} = e^{-\frac{1}{4}\frac{\partial^2}{\partial y''^2}} e^{-(y''-y')^2}$$

$$= \frac{1}{\sqrt{4\pi}} e^{-\frac{1}{4}\frac{\partial^2}{\partial y''^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{4}k^2 + i(y''-y')k}\, dk$$

$$= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} e^{i(y''-y')k}\, dk = \sqrt{\pi}\, \delta(y'-y'') \qquad (9.153)$$

Hence

$$\sum_{n'} \langle \rho'' | n' \rangle \langle n' | \rho' \rangle = \Delta\omega \sqrt{\frac{m\omega}{\hbar}}\, e^{-\frac{1}{2}(y'^2 - y''^2)} \delta(y'-y'')$$

$$= \Delta\omega \sqrt{\frac{m\omega}{\hbar}}\, \delta\left(\sqrt{\frac{m\omega}{\hbar}}(\rho'-\rho'')\right) = \Delta\omega\, \delta(\rho'-\rho'')$$

$$= \delta_{\rho'\rho''}, \qquad (9.154)$$

and the completeness is proved.

c) Systems with angular momentum

In chapter 4 we considered systems for which a total angular momentum vector J can be defined. The components of J were shown to satisfy the Poisson bracket relation (see (4.11)).

$$(J_i, J_j) = \varepsilon_{ijk} J_k \qquad (9.155)$$

In the quantum theory of these systems there will be analogous operator quantities $\mathcal{J}_i$ satisfying the commutation relation

$$[\mathcal{J}_i, \mathcal{J}_j] = i\hbar \varepsilon_{ijk} \mathcal{J}_k . \qquad (9.156)$$

We shall now study the properties of these operators.

In chapter 4 we saw that if A is any vector or tensor quantity composed out of the canonical variables, the variation in A produced by a rotation of the Cartesian coordinates through an angle $-\delta\alpha$ is given in classical theory by (see(4.9))

$$\delta A = (A, \mathcal{J}_i)\, \delta\alpha_i . \qquad (9.157)$$

We hall expect an analogous relation to hold in quantum theory.

$$\delta A = \frac{1}{i\hbar} [A, \mathcal{J}_i]\, \delta\alpha_i . \qquad (9.158)$$

In particular, if A is a scalar then δ A = 0, and

$$[A, \mathcal{J}_i] = 0 . \qquad (9.159)$$

Of particular interest is the square of the total angular momentum

$$\mathcal{J}^2 = \mathcal{J}_i \mathcal{J}_i . \qquad (9.160)$$

Since J^2 is a scalar we should expect it to commute with the J_i. This may be verified directly

$$\begin{aligned} [\mathcal{J}^2, \mathcal{J}_i] &= [\mathcal{J}_j \mathcal{J}_j, \mathcal{J}_i] \\ &= [\mathcal{J}_j, \mathcal{J}_i]\mathcal{J}_j + \mathcal{J}_j [\mathcal{J}_j, \mathcal{J}_i] \\ &= i\hbar \varepsilon_{jik} (\mathcal{J}_k \mathcal{J}_j + \mathcal{J}_j \mathcal{J}_k) = 0 . \end{aligned} \qquad (9.161)$$

Now, the expectation value of J^2 in any state must evidently be always greater than or equal to zero. Hence its eigenvalues must be greater than or equal to zero. It will be of interest to consider the operator $J^2 + 1/4\,\hbar^2$, whose eigenvalues are evidently all greater than or equal to $1/4\,\hbar^2$. This operator has an Hermitian square root whose eigenvalues are all greater than or equal to $1/2\,\hbar$. Let us introduce the operator

$$j = \frac{1}{\hbar}\sqrt{J^2 + \frac{1}{4}\hbar^2} - \frac{1}{2}. \qquad (9.162)$$

j is an Hermitian operator whose eigenvalues are all greater than or equal to zero. Since j is a function of J^2 it commutes with each of the J_i. In this section we shall not be considering any operators other than the J_i. Hence, one having selected an eigenvalue for j, we may, in our subsequent work, regard j as being simply a number j', equal to that eigenvalue. Of course, we have not yet determined the possible eigenvalues of j. That is what we shall do in the following paragraphs.

Let us now introduce also the following operators :

$$j_+ = \frac{1}{\hbar}(J_1 + iJ_2), \quad j_- = \frac{1}{\hbar}(J_1 - iJ_2), \qquad (9.163)$$

$$m = \frac{1}{\hbar}J_3. \qquad (9.164)$$

We may write

$$\begin{aligned} j_+ j_- &= \frac{1}{\hbar^2}(J_1 + iJ_2)(J_1 - iJ_2) \\ &= \frac{1}{\hbar^2}(J_1^2 + J_2^2 - i[J_1, J_2]) \\ &= \frac{1}{\hbar^2}(J^2 - J_3^2 + \hbar J_3). \end{aligned} \qquad (9.165)$$

Making use of (9.164) and the fact that (cf. (5.37))

$$\mathbf{J}^2 = \hbar^2\left(j+\frac{1}{2}\right)^2 - \frac{1}{4}\hbar^2 = \hbar^2 j(j+1), \tag{9.166}$$

we may write (9.165) in the form

$$J_+J_- = \hbar^2 j(j+1) - \hbar^2 m^2 + \hbar^2 m. \tag{9.167}$$

Similarly

$$J_-J_+ = \hbar^2 j(j+1) - \hbar^2 m^2 - \hbar^2 m. \tag{9.168}$$

Hence

$$[J_+, J_-] = 2\hbar^2 m. \tag{9.169}$$

Also

$$[m, J_+] = \frac{1}{\hbar^2}[J_3, J_1 + iJ_2]$$

$$= \frac{1}{\hbar^2}(iJ_2 + J_1) = J_+ \tag{9.170}$$

and similarly

$$[m, J_-] = -J_- \tag{9.171}$$

Let $|m'\rangle$ be an eigenvector of m corresponding to the eigenvalue m'. Then, using (9.167), we may write

$$\hbar^2 j(j+1) - \hbar^2 m'(m'-1) = \langle m'|J_+(J_- + \hbar^2 m)|m'\rangle$$

$$= (\langle m'|J_+)(J_-|m'\rangle) \geq 0, \tag{9.172}$$

from which we may infer

$$-j' \leq m' \leq j'+1 . \tag{9.173}$$

Similarly, using (9.168), we have

$$j'(j'+1) - m'(m'+1) = \langle m' | j_- | j_+ | m' \rangle \geq 0 \tag{9.174}$$

which implies

$$-j'-1 \leq m' \leq j' . \tag{9.175}$$

Equations (9.173) and (9.175) combined give

$$-j' \leq m' \leq j' . \tag{9.176}$$

The equality sign can hold in (9.172) only if $j_- | m' \rangle = 0$. If $j_- | m' \rangle \neq 0$, then, using (9.171), we obtain

$$m j_- | m' \rangle = ([m, j_-] + j_- m) | m' \rangle = (m'-1) j_- | m' \rangle \tag{9.177}$$

which says that $j_- | m' \rangle$ is an eigenvector of m corresponding to the eigenvalue m' - 1. Similarly, if $j_-^2 | m' \rangle \neq 0$, it is an eigenvector of m corresponding to the eigenvalue m' - 2. Continuing in this manner, repeatedly applying the operator j_- , we see that m must have a series of eigenvalues m', m' - 1, m' - 2, ... Unless the equality sign in (9/172) holds at some point in this procedure, the eigenvalues of m will eventually become less that - j', in contradiction with (9.173). The lowest eigenvalue of m is therefore that for which the equality sign in (9.172) holds, namely - j'. We have

$$j_- | -j' \rangle = 0 . \tag{9.178}$$

The equality sign in (9.174), on the other hand can hold only if $j_+ | m' \rangle = 0$. If $j_+ | m' \rangle \neq 0$, then, using (9.170), we obtain

$$m\,j_+|m'\rangle = ([m,j_+]+j_+m)|m'\rangle = (m'+1)j_+|m'\rangle \quad (9.179)$$

which says that $j_+|m'\rangle$ is an eigenvector of m corresponding to the eigenvalue $m'+1$. By repeated applications of the operator j_+ we obtain eigenvectors corresponding to the series of eigenvalues m', $m'+1$, $m'+2$, ... This series must evidently terminate in j' for which value the equal by sign in (9.174) holds:

$$j_+|j'\rangle = 0 \quad (9.180)$$

We now see that the eigenvalues of m have integral spacing and are given by (cf. (5.29))

$$m' = -j', -j'+1, -j'+2, \ldots j'-2, j'-1, j' \quad (9.181)$$

(9.181) now tells us something about the eigenvalues of j. We see that j' can take on only the values

$$j' = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, \ldots \quad (9.182)$$

That is, $2j'$ must be an integer. Let us now indicate the possible variability in j' by writing the eigenvectors in the form $|j', m'\rangle$. This is permissible since j and m commute with one another and are hence simultaneously observable. Let us write

$$\langle j', m'|j_+ j_-|j', m'\rangle = \langle j', m'|j^2 - m^2 + m|j', m'\rangle$$

$$= [j'(j'+1) - m'(m'-1)]\langle j', m'|j', m'\rangle = j'(j'+1) - m'(m'-1) \quad (9.183)$$

For normalization we see that we must have

$$|j', m'-1\rangle = \frac{1}{\sqrt{j'(j'+1) - m'(m'-1)}}\, j_-|j', m'\rangle \quad (9.184)$$

Similarly, from

$$\langle j', m'-1 | j_- j_+ | j', m'-1 \rangle = j'(j'+1) - m'(m'-1) \quad (9.185)$$

we obtain

$$|j', m'\rangle = \frac{1}{\sqrt{j'(j'+1) - m'(m'-1)}}\, j_+ |j', m'-1\rangle \quad (9.186)$$

The matrix elements of the operators j_+ and j_- are seen to be

$$\langle j', m' | j_+ | j', m'-1 \rangle = \sqrt{j'(j'+1) - m'(m'-1)} \quad (9.187)$$

$$\langle j', m'-1 | j_- | j', m' \rangle = \sqrt{j'(j'+1) - m'(m'-1)} \quad (9.188)$$

all other matrix elements vanishing.

From (9.166) we see that the eigenvalues of the square of the angular momentum vector are

$$J'^2 = j'(j'+1)\hbar^2 \quad (91.89)$$

with the j's given by (9.182). Similarly, from (9.164), we see that the eigenvalues of the third component of the angular momentum vector are

$$J'_3 = m'\hbar \quad (9.190)$$

with the m's given by (9.181). From symmetry it is evident that these are also the eigenvalues of any component of J.

We see from equations (9.184) and (9.186) that if any function of the operators j_+, j_-, m is applied to any linear combination of the vectors $|j', m'\rangle$, the result will again be a linear combination of the vectors. Since all the J_i are expressible in terms of j_+, j_-, m, this is also true of any function of the J_i. Hence, so far as the J_i are concerned, the vectors $|j', m'\rangle$ form a complete set, the ope-

rators j and m forming together a complete set of commuting proper real operators.

Using (9.187, 188), and the equations

$$J_1 = \frac{1}{2}\hbar(J_+ + J_-), \tag{9.191}$$

$$J_2 = -\frac{i}{2}\hbar(J_+ - J_-), \tag{9.192}$$

$$J_3 = \hbar m, \tag{9.193}$$

we may readily determine the matrix elements of the components of J.

$$\langle j', m'-1|J_1|j, m\rangle = \langle j', m|J_1|j, m'-1\rangle$$

$$= \frac{1}{2}\hbar\sqrt{(j'+1)-m'(m'-1)} = \frac{1}{2}\hbar\sqrt{(j+m')(j-m'+1)}, \tag{9.194}$$

$$\langle j', m'-1|J_2|j, m'\rangle = -\langle j', m|J_2|j, m'-1\rangle$$

$$= \frac{i}{2}\hbar\sqrt{(j'+m')-m'(m'-1)} = \frac{i}{2}\hbar\sqrt{(j+m')(j-m'+1)}, \tag{9.195}$$

$$\langle j', m|J_3|j, m'\rangle = \hbar m', \tag{9.196}$$

all other matrix elements vanishing.

If the total angular momentum of the system is separable into an orbital and a spin part which satisfy the Poisson bracket relations (4.72), (4.73) and (4.36), then these parts may be treated separately in just the same manner as the total angular momentum has been treated above. The quantum operators L and S satisfy the Poisson bracket relations

$$[L_i, L_j] = i\hbar\,\varepsilon_{ijk}L_k, \tag{9.197}$$

$$[S_i, S_j] = i\hbar \epsilon_{ijk} S_k \qquad (9.198)$$

$$[L_i, S_j] = 0, \qquad (9.199)$$

and are, of course, related to the total angular momentum operator J by

$$J = L + S \qquad (9.200)$$

The quantum numbers used to fix the magnitudes of the orbital and spin angular momenta are customarily denoted by ℓ and S respectively. There is no accepted convention for denoting the quantum numbers which fix the components L_3 and S_3. We shall here denote them simply by m_ℓ and m_s respectively. Thus

$$L'^2 = \hbar^2 \ell'(\ell'+1), \qquad (9.201)$$

$$L'_3 = \hbar m_\ell, \qquad (9.202)$$

$$S'^2 = \hbar^2 s'(s'+1) \qquad (9.203)$$

$$S'_3 = \hbar m_s. \qquad (9.204)$$

As we have seen in chapter 4, the orbital angular momentum is always describable in terms of the canonical coordinates and their conjugate momenta. The spin angular momenta, on the other hand, in the case that they refer to point particles which have been conceptually abstracted from rotating rigid bodies, need not be describable in terms of canonically conjugated variables. They therefore represent a special case. We shall consider them presently, but we shall first turn our attention to orbital angular momenta.

Let ρ^i denote the coordinates of the system, which, together with their conjugate momenta, are sufficient to define the orbital angular momentum of the system. Suppose that none of the spin variables of the system are included among the ρ^i. Now let the coordinate system be rotated through an angle $-\delta\alpha$. Then, by eq. (9.158), the new coordinates, which we may denote by ρ_N^i will be given in terms of the old coordinates ρ_0^i by

$$\rho_N^i = \rho_0^i + \frac{1}{i\hbar}[\rho_0^i, J_\alpha]\delta\alpha_\alpha \qquad (9.205)$$

Since the ρ_0^i do not include the spin variables, the spin part of J commutes with the ρ_0^i, and eq. (9.205) reduces to

$$\rho_N^i = \rho_0^i + \frac{1}{i\hbar}[\rho_0^i, L_\alpha]\delta\alpha_\alpha$$
$$= U^{-1}\rho_0^i U, \qquad (9.206)$$

where

$$U = 1 + \frac{1}{i\hbar} L\cdot\delta\alpha = e^{-\frac{i}{\hbar}L\cdot\delta\alpha} \qquad (9.207)$$

Suppose the infinitesimal pseudovector $\delta\alpha$ is taken parallel to the x_3-axis and equal in magnitude to $\delta\varphi$. Then U becomes

$$U = e^{-\frac{i}{\hbar}L_3\delta\varphi}, \qquad (9.208)$$

and, for an arbitrary rotation $-\varphi$ about the x_3-axis,

$$U = e^{-\frac{i}{\hbar}L_3\varphi} \qquad (9.209)$$

Let us now introduce the configuration space representation. We shall use the symbol ρ' to denote the coordinate eigenvalues, whether they be of the old or the new coordinates. This means that we must distinguish between the eigenvectors $|\rho'\rangle_0$ of the

old coordinates and the eigenvectors $|\rho'\rangle_N$ of the new. We may write

$$\rho'^i |\rho'\rangle_0 = \rho_0^i |\rho'\rangle_0 , \tag{9.210}$$

$$\rho'^i |\rho'\rangle_N = \rho_N^i |\rho'\rangle_N$$
$$= U^{-1} \rho_0^i U |\rho'\rangle_N . \tag{9.211}$$

Multiplying (9.211) on the left by U, we obtain

$$\rho_0^i U |\rho'\rangle_N = \rho'^i U |\rho'\rangle_N , \tag{9.212}$$

which tells us that the $|\rho'\rangle_0$ and $|\rho'\rangle_N$ are related by

$$|\rho'\rangle_0 = U |\rho'\rangle_N , \tag{9.213}$$

or $$|\rho'\rangle_N = U^{-1} |\rho'\rangle_0 , \tag{9.214}$$

and, since U is unitary

$${}_N\langle\rho'| = {}_0\langle\rho'| U . \tag{9.215}$$

If the quantum state of the system is $|\ \rangle$, then ${}_0\langle\rho'|\ \rangle$ and ${}_N\langle\rho'|\ \rangle$ define respectively the values of the wave function at the point in the old coordinate system and the point in the new coordinate system which have the same coordinate values. Let us take $|\ \rangle$ to be an eigenstate $|\ell', m'_\ell, \gamma'\rangle$ of the operator L_3. Then, using (9.215) and (9.209), we have

$${}_N\langle\rho'|\ell', m'_\ell, \gamma'\rangle = {}_0\langle\rho'| e^{-\frac{i}{\hbar} L_3 \varphi} |\ell', m'_\ell, \gamma'\rangle$$
$$= e^{-i m'_\ell \varphi} {}_0\langle\rho'|\ell', m'_\ell, \gamma'\rangle . \tag{9.216}$$

Now, when $\varphi = 2\pi$, the coordinates suffer one complete rotation and become identical with the original coordinates again. The wave functions must return to their original values, and hence we must have

$$e^{-2\pi i m'_z} = 1. \tag{9.217}$$

Equation (9.217) can be satisfied, however, only for integral values of m'_z. By equation (9.181), inserting l' for j' and m'_z for m', this m and what l' can have only integral values (cf. (5.28)).

$$l' = 0, 1, 2, \ldots . \tag{9.218}$$

The sign angular momentum quantum numbers, on the other hand, are under no such restriction, provided they refer to point particles whose spin variables do not constitute a complete classically canonical set. The spin quantum numbers S', m'_s may assume half- old-integral as well as integral values. The case $S' = 1/2$ is of special importance, being that which appears to describe the spin properties of electrons, protons, neutrinos, and muons, and we shall consider it now in some detail.

Using eqs. (9.194-196), and substituting $S' = 1/2$ for j' and $m'_s = \mp 1/2$ for m', we obtain for the matrix elements of the components of the spin angular momentum vector,

$$\left\langle \tfrac{1}{2}, -\tfrac{1}{2} \middle| S_1 \middle| \tfrac{1}{2}, \tfrac{1}{2} \right\rangle = \left\langle \tfrac{1}{2}, \tfrac{1}{2} \middle| S_1 \middle| \tfrac{1}{2}, -\tfrac{1}{2} \right\rangle = \frac{\hbar}{2} \tag{9.219}$$

$$\left\langle \tfrac{1}{2}, -\tfrac{1}{2} \middle| S_2 \middle| \tfrac{1}{2}, \tfrac{1}{2} \right\rangle = -\left\langle \tfrac{1}{2}, \tfrac{1}{2} \middle| S_2 \middle| \tfrac{1}{2}, -\tfrac{1}{2} \right\rangle = \frac{i\hbar}{2} \tag{9.220}$$

$$\left\langle \tfrac{1}{2}, \pm\tfrac{1}{2} \middle| S_3 \middle| \tfrac{1}{2}, \pm\tfrac{1}{2} \right\rangle = \pm \frac{1}{2}\hbar, \tag{9.221}$$

The matrix representations of the S may therefore be expressed in the form

$$S_i = \frac{1}{2}\hbar\sigma_i, \tag{9.222}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{9.223}$$

The σ_i are known as the Pauli spin matrices. They evidently satisfy the equation

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k. \tag{9.224}$$

By direct computation it is easy to verify that they also satisfy the equation

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}. \tag{9.225}$$

To construct the wave function corresponding to a typical system possessing spin angular momenta it is necessary to introduce labels referring to the spin variables in addition to those referring to the coordinates. If, as is the typical situation, the square of the spin angular momentum vector commutes with the Hamiltonian function of the system then the quantum number S' remains permanently constant and may be dropped from the formalism, its value always being implicitly understood. The basis vectors may then be labeled as $|\rho', m'_s\rangle$ and the wave function of the system written as

$$\psi_{m'_s}(\rho', t) = \Delta\omega^{-1/2}\langle\rho', m'_s|t\rangle. \tag{9.226}$$

It is frequently the custom to combine the spin components of the wave function into a column vector, thus :

$$\psi(\rho;t) = \begin{pmatrix} \psi_{s'}(\rho;t) \\ \psi_{s'-1}(\rho;t) \\ \vdots \\ \psi_{-s'+1}(\rho;t) \\ \psi_{-s'}(\rho;t) \end{pmatrix} \qquad (9.227)$$

The spin operators are then conveniently taken in their matrix forms as defined by eqs. (9.194-196). When this done, however, it is necessary to treat orthogonal coordinate transformations in a special manner. Consider an infinitesimal rotation of the Cartesian coordinates x_i of the form

$$x_i' = x_i + \varepsilon_{ij} x_j \qquad (9.228)$$

where ε_{ij} is an infinitesimal antisymmetric in the indices i and j. Normally we would say that the spin angular momentum vector transforms under this transformation according to

$$S_i' = S_i + \varepsilon_{ij} S_j , \qquad (9.229)$$

so that

$$\overline{S_i'} = \langle t | S_i | t \rangle = \int \psi^*(\rho;t) S_i \psi(\rho;t) d\omega'$$
$$= \int \psi^*(\rho;t)(S_i + \varepsilon_{ij} S_j)\psi(\rho;t) d\omega' \qquad (9.230)$$

However, since we have found it convenient to introduce a particular set of matrices for the S_i, we do not wish these matrices to be altered by a coordinate transformation. We can instead let the wave function ψ assume the burden of transformation. The transformation must be unitary, and hence we write

$$\psi' = U\psi \quad , \quad \psi'^{*} = \psi U^{*} \tag{9.231}$$

where, since the transformation is infinitesimal, U has the form

$$U = 1 + iA,$$

A being Hermitian with vanishing trace. We then write

$$\begin{aligned} \overline{S_i'} &= \int \psi'^{*} S_i \psi' \, d\omega' \\ &= \int \psi^{*} S_i \psi \, d\omega' + i \int \psi^{*} [S_i, A] \psi \, d\omega', \end{aligned} \tag{9.232}$$

We must evidently have

$$\varepsilon_{ij} S_j = i [S_i, A]. \tag{9.233}$$

The solution of (9.233), as may be readily verified by direct substitution, is

$$\begin{aligned} A &= -\frac{i}{2\hbar^2} \varepsilon_{ij} [S_i, S_j] \\ &= \frac{1}{2\hbar} \varepsilon_{ij} \varepsilon_{ijk} S_k. \end{aligned} \tag{9.234}$$

Thus the variation in ψ under an infinitesimal rotation ε_{ij} is given by

$$\delta\psi = \frac{1}{2\hbar^2}\epsilon_{ij}[S_i, S_j]\psi. \qquad (9.235)$$

A system for which the spin angular momentum quantum number is 0, 1/2, 1, ... etc. is said to be a system of spin 0, spin 1/2; spin 1, ... etc. For a system of spin 1/2, equation (9.235) may be written in terms of the Pauli matrices (9.223) as

$$\delta\psi = \frac{1}{8}\epsilon_{ij}[\sigma_i, \sigma_j]\psi. \qquad (9.236)$$

The column vector ψ has in this case two components and is called a spinor. Evidently, from the results of (9.216), a spinor changes sign under rotation of the coordinates through an angle 2π. More generally, the wave function possessed by a system of spin - S' has $2S' + 1$ components and is multiplied by the factor $(-1)^{2S'}$ under a rotation of the coordinates through an angle 2π.

Problem XLVIII : Show that the spin matrices of a system of spin -1 are given by

$$S_1 = \hbar\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad S_2 = \hbar\begin{pmatrix} 0 & -\frac{i}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ 0 & \frac{i}{\sqrt{2}} & 0 \end{pmatrix}, \quad S_3 = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Show that under the unitary transformation $\varphi = U\psi$, where

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & -1 & 0 \end{pmatrix},$$

the transformed spin matrices $S_i' = U S_i U^{-1}$ are given by

$$S_1' = \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad S_2' = \hbar \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad S_3' = \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Show that the new wave function transforms like an ordinary 3-vector un er coordinates rotations, i.e;

$$\varphi_i' = \varphi_i + \epsilon_{ij} \varphi_j ,$$

and that the expectation value of the spin angular momentum vector may be expressed in the form

$$\bar{S} = -i\hbar \int \varphi^* \times \varphi \, d\omega' .$$

Sometimes we know the total angular momentum quantum numbers j' and m' of a system when it is in a certain state, together with the orbital and spin quantum numbers ℓ' and S' we wish to know the probability of finding it in the various orbital and spin states $|\ell', S', m_\ell', m_S'\rangle$. This problem is a special case of the following more general problem.

Consider a system formed from the combination of two separate systems having total spin angular momentum vectors J_1 and J_2 respectively. The total angular momentum vector of the combined system is then (cf. (9.200))

$$J = J_1 + J_2 . \qquad (9.237)$$

We have the commutation relations

$$\left.\begin{aligned} &[J_{1i}, J_{1j}] = i\hbar\,\varepsilon_{ijk} J_{1k}, \quad [J_{2i}, J_{2j}] = i\hbar\,\varepsilon_{ijk} J_{2k} \\ &[J_1, J_2] = 0, \quad [J_i, J_j] = i\hbar\,\varepsilon_{ijk} J_k \\ &[J_1^2, J_1] = 0, \quad [J_2^2, J_2] = 0, \quad [J^2, J] = 0 \end{aligned}\right\} \tag{9.238}$$

and we may write

$$J^2 = J_1^2 + J_2^2 + 2 J_1 \cdot J_2 \tag{9.239}$$

and

$$\left.\begin{aligned} &[J, J_1^2] = [J_1 + J_2, J_1^2] = 0 \\ &[J, J_2^2] = [J_1 + J_2, J_2^2] = 0 \\ &[J^2, J_1^2] = 0, \quad [J^2, J_2^2] = 0 \end{aligned}\right\} \tag{9.240}$$

Evidently we may choose as quantum numbers either the set j_1, j_2, j, m, or the set j_1, j_2, m_1, m_2. Our problem is to find the transformation coefficients $\langle j_1, j_2, m_1, m_2 | j_1, j_2, j, m \rangle$.

We shall drop the quantum numbers j_1, j_2, as these will be assumed to be constant throughout the remaining discussion, and we shall write the transformation coefficients simply in the form $\langle m_1, m_2 | j, m \rangle$. We shall also introduce the operators j, m, j_+, j_-, m_1, m_2, etc. as defined by eqs. (9.162-164). We may then write

$$m' | j, m' \rangle = m | j, m' \rangle = (m_1 + m_2) | j, m' \rangle$$

$$= \sum_{m'_1, m'_2} |m'_1, m'_2\rangle \langle m'_1, m'_2 | m_1 + m_2 | j', m'\rangle$$

$$= \sum_{m'_1, m'_2} (m'_1 + m'_2) |m'_1, m'_2\rangle \langle m'_1, m'_2 | j', m'\rangle . \tag{9.241}$$

Multiplying on the left by $\langle m''_1 , m''_2 |$, we obtain

$$m' \langle m''_1, m''_2 | j', m'\rangle = (m''_1 + m''_2) \langle m''_1, m''_2 | j', m'\rangle , \tag{9.242}$$

which implies

$$\langle m'_1, m'_2 | j', m'\rangle = 0 \quad \text{for } m'_1 + m'_2 \neq m' . \tag{9.243}$$

Now, for any given j' the basis vectors $|j', m'\rangle$ for all values of m' from $-j'$ to j' can be generated by repeated application of the operators j_+ and j_- via (9.184) and (9.186). The maximum values of m'_1, and m'_2 are j'_1 and j'_2 respectively. Hence, by (9.243), the maximum value of m' is $j'_1 + j'_2$. This must therefore also be the maximum value of j'. Evidently we may write

$$\langle j'_1, j'_2 | j'_1 + j'_2, j'_1 + j'_2 \rangle = 1 , \tag{9.244}$$

or

$$|j', m'\rangle \Big]_{\substack{j' = j'_1 + j'_2 \\ m' = j'_1 + j'_2}} = |m'_1, m'_2\rangle \Big]_{\substack{m'_1 = j'_1 \\ m'_2 = j'_2}} . \tag{9.245}$$

We can obtain the basis vectors $|j'_1 + j'_2, m'\rangle$ for $m' < j'_1 + j'_2$ by repeated application of equation (9.184), using the fact that $j_- = j_{1-} + j_{2-}$.

The basis vector $|j', m'\rangle_{\substack{j'=j'_1+j'_2-1 \\ m'=j'_1+j'_2-1}}$ must be chosen orthogonal to the vector $|j', m'\rangle_{\substack{j'=j'_1+j'_2 \\ m'=j'_1+j'_2-1}}$. This can be done in only one way since there are only two linearly independent vectors $|m'_1, m'_2\rangle$, with $m'_1 + m'_2 = j'_1 + j'_2 - 1$, with which to construct $|j', m'\rangle_{\substack{j'=j'_1+j'_2 \\ m'=j'_1+j'_2-1}}$ and $|j', m'\rangle_{\substack{j'=j'_1+j'_2-1 \\ m'=j'_1+j'_2-1}}$. This may be illustrated by the following table, where we suppose $j'_1 \geq j'_2$ and where $N_{m'}$ is the number of linearly independent vectors $|m'_1, m'_2\rangle$ with $m'_1 + m'_2 = m'$:

m'	m'_1	m'_2	$N_{m'}$
$j'_1 + j'_2$	j'_1	j'_2	1
$j'_1 + j'_2 - 1$	j'_1	$j'_2 - 1$	2
	$j'_1 - 1$	j'_2	
$j'_1 + j'_2 - 2$	j'_1	$j'_2 - 2$	3
	$j'_1 - 1$	$j'_2 - 1$	
	$j'_1 - 2$	j'_2	
. . . . .			
$j'_1 - j'_2$	j'_1	$-j'_2$	$2j'_2 + 1$
	$j'_1 - 1$	$-j'_2 + 1$	
	. . . . .		
	$j'_1 - 2j'_2$	j'_2	
. . . . .			

m'	m'_1	m'_2	$N_{m'}$
$j'_2 - j'_1$	$2j'_2 - j'_1$	$-j'_2$	$2j'_2 + 1$
	$2j'_2 - j'_1 - 1$	$-j'_2 + 1$	
	. . . . .	. . . . .	
	$-j'_1$	j'_2	
	. . . . .	. . . . .	
$-j'_1 - j'_2 + 1$	$-j'_1 + 1$	$-j'_2$	2
	$-j'_1$	$-j'_2 + 1$	
$-j'_1 - j'_2$	$-j'_1$	$-j'_2$	1

The numbers of values which can be assumed by the quantum numbers m'_1 and m'_2 are $2j'_1 + 1$ and $2j'_2 + 1$ respectively. Hence the number of different basis vectors $|m'_1, m'_2\rangle$ is $(2j'_1 + 1)(2j'_2+1)$, which may also be obtained from the table as follows :

$$\sum_{m'=-j'_1-j'_2}^{j'_1+j'_2} N_{m'} = 2(1+2+3+\ldots 2j'_2) + [2(j'_1-j'_2)+1](2j'_2+1)$$

$$= 2j'_2(2j'_2+1) + (2j'_1-2j'_2+1)(2j'_2+1)$$

$$= (2j'_1+1)(2j'_2+1) \qquad (9.246)$$

This should also be the number of different basis vectors $|j', m'\rangle$.

Having obtained the vector $|j', m'\rangle_{m'=j'_1+j'_2-1}$ from $|j', m'\rangle_{m'=j'_1+j'_2}$, orthogonality requirements, we can obtain the basis vectors $|j'_1 + j'_2 - 1, m'\rangle$ for $m' < j'_1 + j'_2 - 1$ by repeated application

of (9.184). Next, the vector $|j', m'\rangle\Big]_{\substack{j'=j'_1+j'_2-2\\ m'=j'_1+j'_2-2}}$ must be chosen orthogonal to $|j', m'\rangle\Big]_{\substack{j'=j'_1+j'_2\\ m'=j'_1+j'_2-2}}$ and $|j', m'\rangle\Big]_{\substack{j'=j'_1+j'_2-1\\ m'=j'_1+j'_2-2}}$. Again this can be done in only one way. Continuing in this manner, we see that all the coefficients $\langle m'_1, m'_2 | j', m' \rangle$ can be obtained. It is evident that the minimum value of j' is $j'_1-j'_2$ and from our method of construction we see that the number of different basis vectors $|j', m'\rangle$ is given by

$$[2(j'_1+j'_2)+1]+[2(j'_1+j'_2-1)+1]+\ldots+[2(j'_1-j'_2)+1]$$
$$=(2j'_1+2j'_2+1)(2j'_2+1)-2(1+2+3+\ldots+2j'_2)$$
$$=(2j'_1+2j'_2+1)(2j'_2+1)-2j'_2(2j'_2+1) \qquad (9.147)$$
$$=(2j'_1+1)(2j'_2+1),$$

which is the expected result.

As an illustration of the above method we shall consider the combination of two spin 1/2 systems. Then $j'_1 = j'_2 = 1/2$ and we have

$$|1,1\rangle = |1/2, 1/2\rangle \qquad (9.248)$$

$$|1,0\rangle = \frac{1}{1\times 2 - 1\times 0} J_- |1/2, 1/2\rangle$$
$$= \frac{1}{\sqrt{2}}(J_{1-} + J_{2-})|1/2, 1/2\rangle$$
$$= \frac{1}{\sqrt{2}}\sqrt{\tfrac{1}{2}\times\tfrac{3}{2} - \tfrac{1}{2}\times\left(-\tfrac{1}{2}\right)}\,|-1/2, 1/2\rangle$$
$$+ \frac{1}{\sqrt{2}}\sqrt{\tfrac{1}{2}\times\tfrac{3}{2} - \tfrac{1}{2}\times\left(-\tfrac{1}{2}\right)}\,|1/2, -1/2\rangle$$
$$= \frac{1}{\sqrt{2}}|-1/2, 1/2\rangle + \frac{1}{\sqrt{2}}|1/2, -1/2\rangle \qquad (9.249)$$

$$|1,-1\rangle = \frac{1}{\sqrt{1\times 2 - 0\times(-1)}}\,\frac{1}{\sqrt{2}}\,J_-\left(\left|-\tfrac{1}{2},\tfrac{1}{2}\right\rangle + \left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle\right)$$

$$= \frac{1}{2}\left[(j_{1-}+j_{2-})\left|-\tfrac{1}{2},\tfrac{1}{2}\right\rangle + (j_{1-}+j_{2-})\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle\right]$$

$$= \frac{1}{2}\left[0 + \sqrt{\tfrac{1}{2}\times\tfrac{3}{2} - \tfrac{1}{2}\times\left(-\tfrac{1}{2}\right)}\left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle + \sqrt{\tfrac{1}{2}\times\tfrac{3}{2} - \tfrac{1}{2}\times\left(-\tfrac{1}{2}\right)}\left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle + 0\right]$$

$$= \left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle. \qquad (9.250)$$

Let $|0,0\rangle = a\,|-1/2,\ 1/2\rangle + b\,|1/2,\ -1/2\rangle$. We must have $|a|^2 + |b|^2 = 1$ and $a + b = 0$. We may evidently choose $a = -b = \frac{1}{\sqrt{2}}$, so that

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left|-\tfrac{1}{2},\tfrac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle \qquad (9.251)$$

The foregoing method may be extended in an obvious fashion to the treatment of systems formed from the combination of three or more systems having angular momentum. In practive, however, calculation by this method is cumbersome and tedious. Since such systems are of widespread importance in physics, much effort has been expended in developping quicker techniques for handling problems involving angular momentum . Particular mention should be made of the very successful use of the methods of group theory. A study of these methods is unfortunately beyond the scope of this work. For an up to date treatment of angular momentum, in which, however, no systematic use is made of the theory of groups; the reader is referred to J. Schwinger, On Angular Momnetum, Nuclear Development Associates, White Plains, New-York, (1952).

Problem XLIX : Show that for the combination of two systems with $j'_1 = 1$, $j'_2 = 1/2$, we have

$$\left|\tfrac{3}{2}, \tfrac{3}{2}\right\rangle = \left|1, \tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}, \tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\left|0, \tfrac{1}{2}\right\rangle + \tfrac{1}{\sqrt{3}}\left|1, -\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}, -\tfrac{1}{2}\right\rangle = \tfrac{1}{\sqrt{3}}\left|-1, \tfrac{1}{2}\right\rangle + \sqrt{\tfrac{2}{3}}\left|0, -\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{3}{2}, -\tfrac{3}{2}\right\rangle = \left|-1, -\tfrac{1}{2}\right\rangle$$

$$\left|\tfrac{1}{2}, \tfrac{1}{2}\right\rangle = \tfrac{1}{\sqrt{3}}\left|0, \tfrac{1}{2}\right\rangle - \sqrt{\tfrac{2}{3}}\left|1, -\tfrac{1}{2}\right\rangle$$

$$\left|1/2, -\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{2}{3}}\left|-1, \tfrac{1}{2}\right\rangle - \tfrac{1}{\sqrt{3}}\left|0, -\tfrac{1}{2}\right\rangle$$

Problem L : Show that for the combination of two systems with $j'_1 = 1$, $j'_2 = 1$

$$|2, 2\rangle = |1, 1\rangle$$

$$|2, 1\rangle = \tfrac{1}{\sqrt{2}}|0, 1\rangle + \tfrac{1}{\sqrt{2}}|1, 0\rangle$$

$$|2, 0\rangle = \tfrac{1}{\sqrt{6}}|-1, 1\rangle + \sqrt{\tfrac{2}{3}}|0, 0\rangle + \tfrac{1}{\sqrt{6}}|1, -1\rangle$$

$$|2, -1\rangle = \tfrac{1}{\sqrt{2}}|-1, 0\rangle + \tfrac{1}{\sqrt{2}}|0, -1\rangle$$

$$|2, -2\rangle = |-1, -1\rangle$$

$$|1, 1\rangle = \tfrac{1}{\sqrt{2}}|0, 1\rangle - \tfrac{1}{\sqrt{2}}|1, 0\rangle$$

$$|1, 0\rangle = \frac{1}{\sqrt{2}}|-1,1\rangle - \frac{1}{\sqrt{2}}|1,-1\rangle$$
$$|1, -1\rangle = \frac{1}{\sqrt{2}}|-1,0\rangle - \frac{1}{\sqrt{2}}|0,-1\rangle$$
$$|0, 0\rangle = \frac{1}{\sqrt{3}}|-1,1\rangle - \frac{1}{\sqrt{3}}|0,0\rangle + \frac{1}{\sqrt{3}}|1,-1\rangle .$$

In chapter 5 we concluded that a charged multiply periodic system radiates electromagnetic energy, in making a transition from an energy level E_m to a level E_n, at a rate which is proportional to the square of the absolute value of the matrix element (see eq. (5.44))

$$\langle E_n | \eta | E_m \rangle \qquad (9.252)$$

where η is the electric dipole moment of the system. Although this conclusion was based on analogy with classical theory, via the Correspondence Principle, we may anticipate its holding in the exact quantum theory of the interaction of matter and radiation. Now, it frequently happens that the Hamiltonian function of the charged multiply periodic system in question has a form which remains invariant under arbitrary rotations of the system (or, conversely, of the Cartesian axes). In this case the total angular momentum vector J is a constant of the motion, and the individual energy levels are degenerate. The quantum numbers j' and m' (or independent combinations of them) must be introduced in order to specify the quantum states completely. We therefore consider matrix elements of the form

$$\langle j'', m'' | \eta | j', m' \rangle \qquad (9.253)$$

The dipole radiative transition $j', m' \longrightarrow j'', m''$ can take place only if the corresponding matrix element does not vanish. By exami-

ning the matrix elements in detail, therefore, we are led to selection rules. A number of selection rules for simple systems were obtained in chapter 5 through analogy with classical theory via the Correspondence Principle. We shall here obtain some selection rules for systems possessing conserved angular momenta, by completely rigorous arguments. These selecion rules will depend only on the algebraic properties of the angular momentum operators and not on any particular characteristics of the systems themselves.

First, observe that if we have an equation of the form

$$\sum_r f_r(j, m)\, \eta_i\, g_r(j, m) = 0 \qquad (9.254)$$

where the f_r, g_r are functions of the operators j and m, we may infer, by taking the matrix elements, that

$$\langle j'', m'' | \eta_i | j', m' \rangle \sum_r f_r(j'', m'')\, g_r(j', m') = 0, \qquad (9.255)$$

which implies that the matrix elemen (9.253) vanishes unless

$$\sum_r f_r(j'', m'')\, g_r(j', m') = 0 \qquad (9.256)$$

Now, since the electric dipole moment is a vector composed out of the dynamical variables of the system, we have

$$[j_i, \eta_k] = i\hbar\, \epsilon_{ijk}\, \eta_k \qquad (9.157)$$

In particular

$$[m, \eta_3] = \frac{1}{\hbar}[j_3, \eta_3] = 0, \qquad (9.158)$$

or

$$m\eta_3 - \eta_3 m = 0. \qquad (9.259)$$

Also

$$[m,[m,\eta_1]] = i[m,\eta_2] = \eta_1, \tag{9.260}$$

$$[m,[m,\eta_2]] = -i[m,\eta_1] = \eta_2, \tag{9.261}$$

or

$$m^2\eta_1 - 2m\eta_1 m + \eta_1 m^2 - \eta_1 = 0, \tag{9.262}$$

$$m^2\eta_2 - 2m\eta_2 m + \eta_2 m^2 - \eta_2 = 0. \tag{9.263}$$

Equations (9.259), (9.262) and (9.263) are of the form (9.254), and hence we may infer that the matrix elements (9.253) of the various components of η vanish unless

$$m'' - m' = 0, \tag{9.264}$$

or

$$\begin{aligned} 0 &= m''^2 - 2m''m' + m'^2 - 1 \\ &= (m'' - m' - 1)(m'' - m' + 1). \end{aligned} \tag{9.265}$$

We have therefore the following selection rules on m (cf. (5.80))

$$m'' = m', \tag{9.266}$$

or

$$m'' = m' \mp 1 . \tag{9.267}$$

Next, we have

$$[j(j+1),\eta_i] = \frac{1}{\hbar^2}[J^2,\eta_i] = \frac{1}{\hbar^2}[J_j J_j,\eta_i] = \frac{1}{\hbar^2}\{J_j,[J_j,\eta_i]\}$$

$$= \frac{i}{\hbar} \varepsilon_{ijk} \{J_i, \eta_k\} = -\frac{i}{\hbar} \varepsilon_{ijk} \{J_i, \eta_k\}, \tag{9.268}$$

and

$$[j(j+1), [j(j+1), \eta_i]] = -\frac{i}{\hbar} \varepsilon_{ijk} \{J_i, [j(j+1), \eta_k]\}$$

$$= -\frac{1}{\hbar^2} \varepsilon_{ijk} \varepsilon_{klm} \{J_i, \{J_l, \eta_m\}\}$$

$$= -\frac{1}{\hbar^2} (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})(J_i J_l \eta_m + J_i \eta_m J_l + J_l \eta_m J_i + \eta_m J_l J_i)$$

$$= -\frac{1}{\hbar^2} (J_i J_i \eta_j + J_i \eta_j J_i + J_l \eta_j J_l + \eta_j J_i J_i - J^2 \eta_j - 2 J_i \eta_j J_i - \eta_j J^2)$$

$$= -\frac{1}{\hbar^2} [2\{J_i \eta_j J_i\} + [J_i, J_j]\eta_i - \eta_i [J_i, J_j] - 2\{J^2 \eta_j\} - J_i [\eta_i, J_j] - [J_i, \eta_j] J_i]$$

$$= 2\{j(j+1), \eta_j\} - \frac{2}{\hbar^2}\{J_i \eta_j J_i\} - \frac{1}{\hbar^2}([\eta_j, [J_i, J_i]] - [J_i, [J_i, \eta_j]])$$

$$= 2\{j(j+1), \eta_j\} - \frac{2}{\hbar^2}\{J_i \eta_j J_i\} - \frac{1}{\hbar} \varepsilon_{ijk} ([J_i, J_j \eta_k] - [J_i, \eta_k J_j]) \tag{9.269}$$

$$= 2\{j(j+1), \eta_j\} - \frac{2}{\hbar^2}\{J_i \eta_j J_i\}$$

or

$$\hbar^2 (j+1)^2 \eta_j - 2\hbar(j+1)\eta_j \hbar(j+1) + \eta_j \hbar^2 (j+1)^2 - 2\hbar(j+1)\eta_j - 2\eta_j \hbar(j+1) = -\frac{2}{\hbar^2}\{J_i \eta_j J_i\} \tag{9.270}$$

Except for the term on the right hand side, this equation is of the form (9.254). Let us therefore examine the quabtity on the right a little more closely. Let us book particularly at the scalar product

$\gamma \cdot \eta$. Since this quantity is a scalar it commutes with the J_i and hence with J^2 and j. That is

$$j(\gamma \cdot \eta) - (\gamma \cdot \eta) j = 0 \qquad (9.271)$$

Equation (9.271) is of the form (9.254), with η_i replaced by $\gamma \cdot \eta$. Hence we may infer that the matrix elements $\langle j'', m'' | \gamma \cdot \eta | j', m' \rangle$ vanish unless

$$j'' - j' = 0 \qquad (9.272)$$

If $j'' \neq j'$ then

$$\begin{aligned} &\langle j'', m'' | [\gamma \cdot \eta, J_i] | j', m' \rangle \\ &= 2 \sum_{j''', m'''} \langle j'', m'' | \gamma_i | j''', m''' \rangle \langle j''', m''' | \gamma \cdot \eta | j', m' \rangle \\ &= 2 \sum_{m'''} \langle j'', m'' | \gamma_i | j'', m''' \rangle \langle j'', m''' | \gamma \cdot \eta | j', m' \rangle \\ &= 0 . \end{aligned} \qquad (9.273)$$

Hence taking those matrix elements of (9.270) for which $j'' \neq j'$, we may infer that the matrix elements $\langle j'', m'' | \eta_i | j', m' \rangle$ vanishes unless $j'' = j'$ or unless

$$\begin{aligned} 0 &= j''^2(j''+1)^2 - 2j''(j''+1)j'(j'+1) + j'^2(j'+1)^2 - 2j''^2(j''+1) - 2j'^2(j'+1) \\ &= [j''(j''+1) - j'(j'+1)]^2 - 4j''^2 + 2j''(j''+1) - 2j'(j'+1) - 4j'' \\ &= [j''(j''+1) - j'(j'+1) + 2j'' + 2][j''(j''+1) - j'(j'+1) - 2j''] \\ &= [(j''+1)^2 - j'(j''+1) + j'(j''+1) - j'(j'+1)][j''^2 - j''j'(j'+1) + j'j''(j'+1) - j'^2(j'+1)] \end{aligned}$$

$$= [(j''+1)+(j'+1)][(j''+1)-(j'+1)][(j''+j'+1)+(j''-j'-1)]$$
$$= (j''+j'+2)(j''+j')(j''-j'+1)(j''-j'-1). \qquad (9.274)$$

Since j', j'' are positive or zero, we cannot have $j'' + j' + 2 = 0$. We can have $j'' + j' = 0$ only if $j' = j'' = 0$. Under these circumstances, however, the J_i are identically zero and $m'=m''=0$. The matrix elements of η_1 and η_2 vanish, and by symmetry $\langle 0, 0 | \eta_3 | 0, 0 \rangle$ also vanishes. Thus, if $j' \neq j''$, the η_i can have non-vanishing elements only if

$$j'' = j' \mp 1. \qquad (9.285)$$

If $j'' = j'$, the matrix form of eq. (9.270) reduces to

$$-4j'(j'+1)\langle j', m'|\eta_3|j', m\rangle = -\sum_{m''} \langle j', m'|J_1|j', m''\rangle \langle j', m''|J_2|j', m\rangle + \sum_{m''} \langle j', m'|J_2|j', m''\rangle \langle j', m''|J_1|j', m\rangle \qquad (9.276)$$

We can have a non vanishing matrix element $\langle j', m'' | \eta_3 | j', m' \rangle$ only if at least one of the matrix elements $\langle j', m'' | J_1 | j', m' \rangle$ $\langle j', m'' | J_2 | j', m' \rangle$ is non-vanishing. This, however, not always the case. For example, if the system consists of one spin particle moving in a central force field (see next section) then

$$\eta_i = \epsilon_{ijk} x_j p_k, \qquad (9.277)$$

where the x_i are the coordinates and the p_i the conjugate momenta, and

$$J_i = \epsilon x_i, \qquad (9.278)$$

so that

$$\mathcal{J}\cdot\eta = e\,\epsilon_{ijk}\,x_j\,p_k\,x_i = e\,\epsilon_{ijk}\,x_i\,x_j\,p_k = 0. \quad (9.279)$$

In this case $j = \ell$, where ℓ is the orbital angular momentum operator, and we have the selection rules (cf. 5.81))

$$\ell'' = \ell' \pm 1. \quad (9.280)$$

Only for more general systems can we have the additional selection rule

$$j'' = j' \quad \text{for } j' \neq 0. \quad (9.281)$$

d) The particle moving in a central force field.

It will be instructive to carry out the quantization of this system in curvilinear coordinates, namely the spherical coordinates appropriate to the system. The Lagrangian function is given by (1.212), and we see that the metric tensor of the spherical coordinate system is given by

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{pmatrix}, \quad (9.282)$$

where we have chosen the coordinates in the order $r = \rho^1$, $\theta = \rho^2$, $\varphi = \rho^3$.

The only non-vanishing Christoffel symbols are the following :

$$\left.\begin{aligned} \left[12,2\right] &= -\left[22,1\right] = r \\ \left[13,3\right] &= -\left[33,1\right] = r\sin^2\theta \\ \left[23,3\right] &= -\left[33,2\right] = r^2\sin\theta\cos\theta \end{aligned}\right\} \quad (9.283)$$

$$\left.\begin{aligned} \left\{\begin{matrix}2\\12\end{matrix}\right\} = \left\{\begin{matrix}3\\13\end{matrix}\right\} &= \frac{1}{r} \;;\; \left\{\begin{matrix}1\\22\end{matrix}\right\} = -r \\ \left\{\begin{matrix}1\\33\end{matrix}\right\} &= -r\sin^2\theta,\; \left\{\begin{matrix}2\\33\end{matrix}\right\} = -\sin\theta\cos\theta,\; \left\{\begin{matrix}3\\23\end{matrix}\right\} = \cot\theta \end{aligned}\right\} \quad (9.284)$$

We therefore have

$$\left\{\begin{matrix}j\\1j\end{matrix}\right\} = \frac{2}{r},\quad \left\{\begin{matrix}j\\2j\end{matrix}\right\} = \cot\theta,\quad \left\{\begin{matrix}j\\3j\end{matrix}\right\} = 0, \quad (9.285)$$

so that the differential operator forms of the conjugate momenta are

$$p_r = -i\hbar\frac{\partial}{\partial r} - i\hbar\frac{1}{r}, \quad (9.286)$$

$$p_\theta = -i\hbar\frac{\partial}{\partial\theta} - \frac{1}{2}i\hbar\cot\theta, \quad (9.287)$$

$$p_\varphi = -i\hbar\frac{\partial}{\partial\varphi}. \quad (9.288)$$

We also have

$$g^{ij}\left\{\begin{matrix}k\\ik\end{matrix}\right\}_j = -\frac{2}{r^2} - \frac{1}{r^2\sin^2\theta}, \quad (9.289)$$

$$g^{ij}\left\{\begin{matrix}1\\ij\end{matrix}\right\} = -\frac{2}{r},\; g^{ij}\left\{\begin{matrix}2\\ij\end{matrix}\right\} = -\frac{\cot\theta}{r^2},\; g^{ij}\left\{\begin{matrix}3\\ij\end{matrix}\right\} = 0, \quad (9.290)$$

$$g^{ij}\left\{\begin{matrix}k\\ij\end{matrix}\right\}\left\{\begin{matrix}l\\kl\end{matrix}\right\} = -\frac{4}{r^2} - \frac{\cot^2\theta}{r^2}, \quad (9.291)$$

$$g^{ij}\left\{\begin{matrix}k\\ik\end{matrix}\right\}\left\{\begin{matrix}l\\jl\end{matrix}\right\} = \frac{4}{r^2} + \frac{\cot^2\theta}{r^2}. \quad (9.292)$$

Hence the function Q of eq. (8.143) becomes

$$Q = \frac{1}{2} g^{ij}\left[\left\{\begin{matrix} k \\ ik \end{matrix}\right\}_{,j} - \left\{\begin{matrix} k \\ ij \end{matrix}\right\}\left\{\begin{matrix} l \\ kl \end{matrix}\right\} - \frac{1}{2}\left\{\begin{matrix} k \\ ik \end{matrix}\right\}\left\{\begin{matrix} l \\ jl \end{matrix}\right\}\right]$$

$$= \frac{1}{r^2 \sin^2\theta}\left(\frac{1}{4}\cos^2\theta - \frac{1}{2}\right), \qquad (9.293)$$

and we may write the Hamiltonian operator in the form (see (1.216))

$$H = \frac{1}{2m}\left[p_r^2 + p_\theta \frac{1}{r^2} p_\theta + p_\varphi \frac{1}{r^2 \sin^2\theta} p_\varphi\right]$$

$$+ \frac{\hbar^2}{2m} \frac{1}{r^2 \sin^2\theta}\left(\frac{1}{4}\cos^2\theta - \frac{1}{2}\right) + V(r). \qquad (9.294)$$

We shall actually find it more convenient for many purposes to leave the Hamiltonian function in Cartesian form

$$H = \frac{1}{2m} p^2 + V(r) \qquad (9.295)$$

p is the momentum vector. Its components are the momenta conjugate to the Cartesian coordinates

$$\begin{aligned} x_1 &= r \sin\theta \cos\varphi \\ x_2 &= r \sin\theta \sin\varphi \\ x_3 &= r \cos\theta \end{aligned} \qquad (9.296)$$

It is of interest to express the radial momentum p_r in Cartesian form. Using (9.294), we have, in the Heisenberg representation.

$$\dot{r} = \frac{1}{i\hbar}[r, H] = \frac{1}{m} p_r \qquad (9.297)$$

But, using (9.295), we get

$$p_r = m\dot{r} = \frac{m}{i\hbar}[r, H] = \frac{1}{2i\hbar}\{[r, p_i], p_i\}$$
$$= \frac{1}{r}\left\{\frac{x_i}{r}, p_i\right\} \qquad (9.298)$$

which has the useful corallary

$$\{r, p_r\} = \frac{1}{2}\left\{r, \left\{\frac{x_i}{r}, p_i\right\}\right\} = \{x_i, p_i\}. \qquad (9.299)$$

p_r may also be written in the form

$$p_r = \frac{x_i}{r} p_i - \frac{i\hbar}{2}\left[\frac{x_i}{r}, p_i\right]$$
$$= \frac{x_i}{r} p_i - \frac{i\hbar}{2}\frac{(3-1)}{r} = \frac{x_i}{r} p_i - \frac{i\hbar}{r} \qquad (9.300)$$
$$= p_i \frac{x_i}{r} + \frac{i\hbar}{r}, \qquad (9.301)$$

the last form being obtainable from the preceding one by taking the Hermitian adjoint.

Now the orbital angular momentum of the particle is given by

$$L_i = \varepsilon_{ijk} x_j p_k \qquad (9.302)$$

and satisfies the following equations

$$L_i x_i = x_i L_i = 0, \quad \text{(see (9.279))} \qquad (9.303)$$
$$L_i p_i = p_i L_i = 0, \qquad (9.304)$$

$$L^2 = \varepsilon_{ijk}\varepsilon_{ilm} x_j p_k x_l p_m$$
$$= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) x_j p_k x_l p_m = x_j p_k x_j p_k - x_j p_k x_k p_j$$

$$= r^2 p^2 + \cdots \varepsilon_{ijk}\left[p_k x_j\right] p_i - p_k x_k x_j p_j - \left[x_j, p_k x_k\right] p_j$$

$$= r^2 p^2 - i\hbar\, r\cdot p - (p_k x_k - i\hbar)(r\cdot p + i\hbar) - i\hbar\, r\cdot p$$

$$= r^2 p^2 - p_k r^2 p_k + i\hbar\left[r, p_k\right] - \hbar^2 - 2i\hbar\, r\cdot p$$

$$= r^2\left(p^2 - p_k^2\right) - \left[p_k, r^2\right] p_k - 2\hbar^2 - i\hbar\left\{x_i, p_i\right\} - i\hbar\left[x_i, p_i\right]$$

$$= r^2\left(p^2 - p_k^2\right) + 2i\hbar\, r\cdot p_k - i\hbar\left\{r, p_k\right\} + \hbar^2$$

$$= r^2\left(p^2 - p_k^2\right) + i\hbar\left[r, p_k\right] + \hbar^2$$

$$= r^2\left(p^2 - p_r^2\right) \qquad (9.305)$$

Since r^2 is a scalar it commutes with L^2. Hence, multiplying (9.304) on the left by $\frac{1}{r^2}$, we find we can write (9.295) in the form

$$H = \frac{1}{2m}\left(p_r^2 + \frac{L^2}{r^2}\right) + V(r) \qquad (9.306)$$

Comparison of (9.305) with (9.294) leads us to make the identification (cf. (1.223)

$$L^2 = p_\theta^2 + \frac{1}{\sin^2\theta}\, p_\varphi^2 + \frac{\hbar^2}{\sin^2\theta}\left(\frac{1}{4}\cos^2\theta - \frac{1}{2}\right) \qquad (9.307)$$

Using eqs. (9.287, 288), we may write this in differential operator form

$$L^2 = -\hbar^2\left[\left(\frac{\partial}{\partial\theta} + \frac{1}{2}\cot\theta\right)\left(\frac{\partial}{\partial\theta} + \frac{1}{2}\cot\theta\right) + \frac{\partial^2}{\partial\varphi^2}\frac{1}{\sin^2\theta}\left(\frac{1}{4}\cos^2\theta - \frac{1}{2}\right)\right]$$

$$= -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} - \frac{1}{2\sin^2\theta} + \frac{1}{4}\cot^2\theta + \frac{\partial^2}{\partial\varphi^2}\frac{1}{\sin^2\theta}\left(\frac{1}{4}\cos^2\theta - \frac{1}{2}\right)\right]$$

$$= -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial\varphi^2}\right]$$

$$= -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial\varphi^2} \right]. \qquad (9.308)$$

An operator whose properties will be very useful to us in later sections is the following

$$\mathcal{B}_i = -\frac{1}{2} \varepsilon_{ijk} \{ p_k, L_k \}. \qquad (9.309)$$

This operator may also be expressed in the following forms

$$\mathcal{B}_i = -\varepsilon_{ijk} p_j L_k + \frac{1}{2} \varepsilon_{ijk} [p_j, L_k]$$

$$= -\varepsilon_{ijk} p_j L_k + \frac{i\hbar}{2} \varepsilon_{ijk} \varepsilon_{jke} p_e$$

$$= -\varepsilon_{ijk} p_j L_k + i\hbar p_i \qquad (9.310)$$

$$= \varepsilon_{ijk} L_j p_k - i\hbar p_i, \qquad (9.311)$$

the last form being obtained from the preceding one by taking the Hermitian adjoint. $\mathcal{B}_i$ satisfies the following equations.

$$\mathcal{B}^2 = (\varepsilon_{ijk} L_j p_k - i\hbar p_i)(-\varepsilon_{ilm} p_e L_m + i\hbar p_i)$$

$$= -(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) L_j p_k p_e L_m + \hbar^2 p^2$$

$$= L_j p^2 L_j + \hbar^2 p^2 = (L^2 + \hbar^2) p^2, \qquad (9.312)$$

$$[\mathcal{B}_i, p_j] = -\frac{1}{2} \varepsilon_{ikl} \{ p_k [L_l, p_j] \}$$

$$= -\frac{i\hbar}{2} \varepsilon_{ikl} \varepsilon_{ljm} \{ p_k, p_m \}$$

$$= -i\hbar(\delta_{ij}\delta_{km} - \delta_{im}\delta_{kj})p_k p_m$$

$$= -i\hbar(\delta_{ij}p^2 - p_i p_j),$$

(9.313)

$$\{B_i, \frac{x_i}{r}\} = (\varepsilon_{ijk}L_j p_k - i\hbar p_i)\frac{x_i}{r} + \frac{x_i}{r}(-\varepsilon_{ijk}p_k L_k + i\hbar p_i)$$

$$= -2\frac{L^2}{r} + i\hbar\left[\frac{x_i}{r}, p_i\right] = -2\frac{L^2}{r} - \hbar^2\frac{(3-1)}{r}$$

$$= -\frac{2}{r}(L^2 + \hbar^2),$$

(9.314)

$$\left[B_i, \frac{x_i}{r}\right] = -\frac{1}{2}\varepsilon_{ikl}\left\{\left[p_k, \frac{x_i}{r}\right], L_l\right\} - \frac{1}{2}\varepsilon_{ikl}\left\{p_k, \left[L_l, \frac{x_i}{r}\right]\right\}$$

$$= \frac{i\hbar}{2}\varepsilon_{ikl}\left\{\frac{\delta_{ik}}{r} - \frac{x_i x_k}{r^3}, L_l\right\} - \frac{i\hbar}{2}\varepsilon_{ikl}\varepsilon_{ljm}\left\{p_k, \frac{x_m}{r}\right\}$$

$$= \frac{i\hbar}{r}\varepsilon_{ijk}L_k - \frac{i\hbar}{2}\varepsilon_{ikl}\varepsilon_{lmn}\left\{\frac{x_i x_k}{r^3}, x_m p_n\right\}$$

$$- \frac{i\hbar}{2}(\delta_{ij}\delta_{kn} - \delta_{in}\delta_{kj})\left\{p_k, \frac{x_n}{r}\right\}$$

$$= \frac{i\hbar}{r}\varepsilon_{ijk}L_k - \frac{i\hbar}{2}(\delta_{im}\delta_{kn} - \delta_{in}\delta_{km})\left\{\frac{x_i x_k x_m}{r^3}, p_n\right\}$$

$$- \frac{i\hbar}{2}\delta_{ij}\left\{p_k, \frac{x_k}{r}\right\} + \frac{i\hbar}{2}\left\{p_j, \frac{x_i}{r}\right\}$$

$$= \frac{i\hbar}{r}\varepsilon_{ijk}L_k - \frac{i\hbar}{2}\left\{\frac{x_i x_j x_k}{r^3}, p_k\right\} - i\hbar\delta_{ij}p_k$$

(9.315)

$$+ \frac{i\hbar}{2}\left\{\frac{x_i}{r}, p_j\right\} + \frac{i\hbar}{2}\left\{\frac{x_j}{r}, p_i\right\}.$$

$$[B_i, \tfrac{\alpha_i}{r}] - [B_j, \tfrac{\alpha_i}{r}] = 2\frac{i\hbar}{r}\varepsilon_{ijk}L_k, \qquad (9.316)$$

$$\begin{aligned} L_i B_i &= \varepsilon_{ijk}L_i L_j p_k - i\hbar L_i p_i \\ &= \frac{1}{2}\varepsilon_{ijk}\{L_i, L_j\}p_k + \frac{1}{2}\varepsilon_{ijk}[L_i, L_j]p_k - i\hbar L_i p_i \\ &= \frac{i\hbar}{2}\varepsilon_{ijk}\varepsilon_{ije}L_e p_k - i\hbar L_i p_i = 0, \end{aligned} \qquad (9.317)$$

$$B_i L_i = 0, \qquad (9.318)$$

$$\begin{aligned} [B_i, B_j] &= -\frac{1}{2}\varepsilon_{jke}\{[B_i, p_k]L_e\} - \frac{1}{2}\varepsilon_{jke}\{p_k, [B_i, L_e]\} \\ &= \frac{i\hbar}{2}\varepsilon_{jke}\{\delta_{ik}p^2 - p_i p_k, L_e\} - \frac{i\hbar}{2}\varepsilon_{jke}\varepsilon_{ilm}\{p_k B_m\} \\ &= \frac{i\hbar^2}{2}\varepsilon_{jke}\{\delta_{ik}p^2 - p_i p_k, L_e\} \\ &\quad + \frac{i\hbar^2}{2}(\delta_{ji}\delta_{km} - \delta_{jm}\delta_{ki})\{p_k, \varepsilon_{mab}L_a p_b - i\hbar p_m\} \\ &= \frac{i\hbar^2}{2}\varepsilon_{jke}\{\delta_{ik}p^2 - p_i p_k, L_e\} \\ &\quad + \frac{i\hbar^2}{2}\delta_{ij}\varepsilon_{kab}(p_k L_a p_b + L_a p_b p_k) + \hbar^2\delta_{ij}p^2 \\ &\quad - \frac{i\hbar^2}{2}\varepsilon_{jab}(p_i L_a p_b + L_a p_b p_i) - \hbar^2 p_i p_j \\ &= c\hbar\varepsilon_{jie}L_e p^2 + \frac{i\hbar^2}{2}\delta_{ij}\varepsilon_{kab}p_k[L_a, p_b] + i\hbar^2\delta_{ij}p^2 \\ &\quad - \frac{i\hbar^2}{2}\varepsilon_{jab}p_i[L_a, p_b] - \hbar^2 p_i p_j = -i\hbar\varepsilon_{ijk}L_k p^2. \end{aligned} \qquad (9.319)$$

Let us return now to a study of the differential forms of various operators in both Cartesian and spherical coordinates. From eqs. (9.296) we obtain the transformation matrix

$$\begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial \varphi} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial \varphi} \\ \frac{\partial x_3}{\partial r} & \frac{\partial x_3}{\partial \theta} & \frac{\partial x_3}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} \quad (9.320)$$

and its inverse

$$\begin{pmatrix} \frac{\partial r}{\partial x_1} & \frac{\partial r}{\partial x_2} & \frac{\partial r}{\partial x_3} \\ \frac{\partial \theta}{\partial x_1} & \frac{\partial \theta}{\partial x_2} & \frac{\partial \theta}{\partial x_3} \\ \frac{\partial \varphi}{\partial x_1} & \frac{\partial \varphi}{\partial x_2} & \frac{\partial \varphi}{\partial x_3} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ \frac{1}{r}\cos\theta\cos\varphi & \frac{1}{r}\cos\theta\sin\varphi & -\frac{1}{r}\sin\theta \\ -\frac{1}{r}\frac{\sin\varphi}{\sin\theta} & \frac{1}{r}\frac{\cos\varphi}{\sin\theta} & 0 \end{pmatrix} \quad (9.321)$$

Hence we may write

$$p_1 = -i\hbar\frac{\partial}{\partial x_1} = -i\hbar\left(\frac{\partial r}{\partial x_1}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x_1}\frac{\partial}{\partial\theta} + \frac{\partial\varphi}{\partial x_1}\frac{\partial}{\partial\varphi}\right)$$

$$= -i\hbar\left(\sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - \frac{1}{r}\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}\right), \quad (9.322)$$

$$p_2 = -i\hbar\frac{\partial}{\partial x_2} = -i\hbar\left(\frac{\partial r}{\partial x_2}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x_2}\frac{\partial}{\partial\theta} + \frac{\partial\varphi}{\partial x_2}\frac{\partial}{\partial\varphi}\right),$$

$$= -i\hbar\left(\sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + \frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}\right), \quad (9.323)$$

$$p_3 = -i\hbar\frac{\partial}{\partial x_3} = -i\hbar\left(\frac{\partial r}{\partial x_3}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x_3}\frac{\partial}{\partial\theta} + \frac{\partial\varphi}{\partial x_3}\frac{\partial}{\partial\varphi}\right),$$

$$= -i\hbar\left(\cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\right), \quad (9.324)$$

$$L_1 = x_2 p_3 - x_3 p_2$$

$$= -i\hbar\left[r\sin\theta\sin\varphi\left(\cos\theta\frac{\partial}{\partial z} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\right)\right.$$

$$\left. - r\cos\theta\left(\sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + \frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}\right)\right]$$

$$= -i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta} + \cos\varphi\cot\theta\frac{\partial}{\partial\varphi}\right), \qquad (9.325)$$

$$L_2 = x_3 p_1 - x_1 p_3$$

$$= -i\hbar\left[r\cos\theta\left(\sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - \frac{1}{r}\frac{\sin\varphi}{\cos\theta}\frac{\partial}{\partial\varphi}\right)\right.$$

$$\left. - r\sin\theta\cos\varphi\left(\cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\right)\right]$$

$$= -i\hbar\left(\cos\varphi\frac{\partial}{\partial\theta} - \sin\varphi\cot\theta\frac{\partial}{\partial\varphi}\right). \qquad (9.326)$$

$$L_3 = x_1 p_2 - x_2 p_1$$

$$= -i\hbar\left[r\sin\theta\cos\varphi\left(\sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + \frac{1}{r}\frac{\cos\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}\right)\right.$$

$$\left. - r\sin\theta\sin\varphi\left(\sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - \frac{1}{r}\frac{\sin\varphi}{\sin\theta}\frac{\partial}{\partial\varphi}\right)\right]$$

$$= -i\hbar\frac{\partial}{\partial\varphi}, \qquad (9.327)$$

and

$$l_+ = \frac{1}{\hbar}(L_1 + iL_2) = e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right), \qquad (9.328)$$

$$l_- = \frac{1}{\hbar}(L_1 - iL_2) = e^{-i\varphi}\left(-\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right), \qquad (9.329)$$

$$m = \frac{1}{\hbar}L_3 = -i\frac{\partial}{\partial\varphi}. \qquad (9.330)$$

Using these forms we may now readily obtain the wave functions corresponding to the eigenvectors of the operators

ℓ and m/ First of all we have

$$\langle m | \ell', m' \rangle = \langle m' | \ell', m' \rangle, \tag{9.331}$$

which, when translated into configuration space language, reads

$$-i \frac{\partial}{\partial \varphi} \langle \theta, \varphi | \ell', m' \rangle = m' \langle \theta, \varphi | \ell', m' \rangle \tag{9.332}$$

which says that $\langle \theta, \varphi | \ell', m' \rangle$ is of the form

$$\langle \theta, \varphi | \ell', m' \rangle = \frac{\Delta\omega^{1/2}}{2\pi} e^{im'\varphi} f_{\ell'}^{m'}(\theta), \tag{9.333}$$

where we have inserted normalization factors in such a way that we insure

$$\int_0^\pi |f_{\ell'}^{m'}(\theta)|^2 \sin\theta \, d\theta = 1 \tag{9.334}$$

$\Delta\omega$, or $d\omega$, is here the volume element of angle space (θ, φ) and has the explicit form

$$d\omega = \sin\theta \, d\theta \, d\varphi. \tag{9.335}$$

From eq. (9.180), we have

$$L_+ | \ell', \ell' \rangle = 0, \tag{9.336}$$

which implies

$$\begin{aligned} 0 &= e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta \frac{\partial}{\partial\varphi}\right) e^{i\ell'\varphi} f_{\ell'}^{\ell'}(\theta) \\ &= e^{i(\ell'+1)\varphi}\left(\frac{\partial}{\partial\theta} - \ell'\cot\theta\right) f_{\ell'}^{\ell'}(\theta), \end{aligned} \tag{9.337}$$

or

$$\frac{d f_{\ell'}^{\ell'}}{f_{\ell'}^{\ell'}} = \ell' \cot\theta \, d\theta, \tag{9.338}$$

which on integration yields

$$\log f^{l'}_{l'} = l' \log \sin\theta + \text{constant} \tag{9.339}$$

or

$$f^{l'}_{l'}(\theta) = A \sin^{l'}\theta \tag{9.340}$$

The constant A must be chosen so as to satisfy condition (9.334)
Thus

$$1 = |A|^2 \int_0^\pi \sin^{2l'+1}\theta \, d\theta$$

$$= |A|^2 \left[-\cos\theta \sin^{2l'}\theta\right] + 2l' |A|^2 \int_0^\pi \sin^{2l'-1}\theta \cos^2\theta \, d\theta$$

$$= \frac{2l'}{2l'+1} |A|^2 \int_0^\pi \sin^{2l'+1}\theta \, d\theta = \ldots$$

$$= \frac{2l'(2l'-2)\ldots 2}{(2l'+1)(2l'-1)\ldots 3} |A|^2 \int_0^\pi \sin\theta \, d\theta$$

$$= 2^{2l'+1} \frac{(l'!)^2}{(2l'+1)!} |A|^2 \tag{9.341}$$

We may take

$$A = \frac{\sqrt{(2l'+1)!}}{2^{l'} l'!} \tag{9.342}$$

Using eq. (9.184), we get, for $m' \geq 0$,

$$|l', m'\rangle = \frac{1}{\sqrt{(l'+m')!(l'-m')!}} L^{l'-m'} |l', l'\rangle$$

$$= \frac{1}{\sqrt{2\ell'(4\ell'-2)(6\ell'-2\cdot 3)\cdots[2(\ell'-m')\ell'\cdot(\ell'-m'-1)(\ell'-m')]}}\, L_-^{\ell'-m'}|\ell',\ell'\rangle$$

$$= \frac{1}{\sqrt{1\cdot 2\cdot 3\cdots(\ell'-m')\times 2\ell'(2\ell'-1)(2\ell'-2)\cdots(\ell'+m'+1)}}\, L_-^{\ell'-m'}|\ell',\ell'\rangle$$

$$= \frac{1}{\sqrt{(2\ell')!}}\sqrt{\frac{(\ell'+m')!}{(\ell'-m')!}}\, L_-^{\ell'-m'}|\ell',\ell'\rangle, \qquad (9.343)$$

or

$$\langle\theta,\varphi|\ell',m'\rangle = \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}}\frac{\sqrt{(2\ell'+1)!}}{2^{\ell'}\ell'!}\frac{1}{\sqrt{(\ell'+m')!}}\sqrt{\frac{(\ell'+m')!}{(\ell'-m')!}}$$

$$\times\left[e^{-i\varphi}\left(-\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\varphi}\right)\right]^{\ell'-m'}e^{im'\varphi}\sin^{\ell'}\theta$$

$$= \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}}\sqrt{\frac{2\ell'+1}{2}}\frac{(-1)^{\ell'-m'}}{2^{\ell'}\ell'!}\sqrt{\frac{(\ell'+m')!}{(\ell'-m')!}}e^{im'\varphi}\left(\frac{\partial}{\partial\theta}+im'\cot\theta\right)$$

$$\times\left(\frac{\partial}{\partial\theta}+(m'+2)\cot\theta\right)\times\cdots\times\left(\frac{\partial}{\partial\theta}+\ell'\cot\theta\right)\sin^{\ell'}\theta \qquad (9.344)$$

Now observe that

$$\frac{\partial}{\partial\theta}+n\cot\theta = (\sin\theta)^{-n}\frac{\partial}{\partial\theta}(\sin\theta)^{n} \qquad (9.345)$$

Hence we may write

$$\langle\theta,\varphi|\ell',m'\rangle = \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}}\sqrt{\frac{2\ell'+1}{2}}\frac{(-1)^{\ell'-m'}}{2^{\ell'}\ell'!}\sqrt{\frac{(\ell'+m')!}{(\ell'-m')!}}e^{im'\varphi}(\sin\theta)^{-m'}$$

$$\cdot\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\right)^{\ell'-m'}(\sin\theta)^{2\ell'} \qquad (9.346)$$

In particular

$$\langle \zeta, \varphi | \ell', 0 \rangle = (-1)^{\ell'} \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}}\, P_{\ell'}(\zeta), \tag{9.347}$$

where

$$\zeta = \cos\theta \quad , \quad d\zeta = -\sin\theta\, d\theta, \tag{9.348}$$

and

$$P_{\ell'}(\zeta) \equiv \frac{(-1)^{\ell'}}{2^{\ell'}\ell'!} \left(\frac{d}{d\zeta}\right)^{\ell'} (1-\zeta^2)^{\ell'}. \tag{9.349}$$

The $P_{\ell'}$ are known as Legendre polynomials. ($P_{\ell'}(\zeta)$ is evidently a polynomial of degree ℓ' in ζ.) These polynomials clearly satisfy the orthogonality relations

$$\int_{-1}^{1} P_{\ell''}(\zeta) P_{\ell'}(\zeta)\, d\zeta = \frac{2}{2\ell'+1} \delta_{\ell'\ell''}. \tag{9.350}$$

Now, making use of eq. (9.186), we may rewrite (9.346) in a different form. First of all, we have, for $m' \geq 0$,

$$|\ell', m'\rangle = \frac{1}{\sqrt{[\ell'(\ell'+1)][(\ell'-1)(\ell'+2)][(\ell'-2)(\ell'+3)]\cdots[(\ell'-m'+1)(\ell'+m')]}}\, L_+^{m'} |\ell', 0\rangle$$

$$= \sqrt{\frac{(\ell'-m')!}{(\ell'+m')!}}\, L_+^{m'} |\ell', 0\rangle.$$

$$= \frac{1}{\sqrt{[\ell'(\ell'+1)][(\ell'-1)(\ell'+2)][(\ell'-2)(\ell'+3)]\cdots[(\ell'-m'+1)(\ell'+m')]}}\, L_+^{m'} |\ell', 0\rangle \tag{9.351}$$

With the explicit form (9.329) this becomes, in configuration space

$$\langle \theta, \varphi | \ell', m' \rangle = (-1)^{\ell'} \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \sqrt{\frac{(\ell'-m')!}{(\ell'+m')!}} \left[e^{i\varphi} \left(\frac{\partial}{\partial\theta} + i\cot\theta \frac{\partial}{\partial\varphi}\right)\right]^{m'} P_{\ell'}(\cos\theta)$$

$$= (-1)^{\ell'} \frac{\Delta\omega}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \sqrt{\frac{(\ell'-m')!}{(\ell'+m')!}}\, e^{i m'\varphi} \left(\frac{\partial}{\partial\theta} - (m'-1)\cot\theta\right)\left(\frac{\partial}{\partial\theta} - (m'-2)\cot\theta\right)\times$$

$$\cdots \times \left(\frac{\partial}{\partial\theta} - \cot\theta\right)\frac{\partial}{\partial\theta} P_{\ell'}(\cos\theta)$$

$$= (-1)^{\ell'} \frac{\Delta\omega}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \sqrt{\frac{(\ell'-m')!}{(\ell'+m')!}}\, e^{i m'\varphi} (\sin\theta)^{m'} \left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\right)^{m'} P_{\ell'}(\cos\theta) \qquad (9.352)$$

or

$$\langle\vartheta,\varphi|\ell',m'\rangle = (-1)^{\ell'+m'} \frac{\Delta\omega}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \sqrt{\frac{(\ell'-m')!}{(\ell'+m')!}}\, e^{i m'\varphi} P_{\ell'}^{m'}(\vartheta), \qquad (9.353)$$

where

$$P_{\ell'}^{m'}(\vartheta) \equiv (1-\vartheta^2)^{\frac{m'}{2}} \left(\frac{d}{d\vartheta}\right)^{m'} P_{\ell'}(\vartheta)$$

$$= \frac{(-1)^{\ell'}}{2^{\ell'}\ell'!} (1-\vartheta^2)^{-\frac{m'}{2}} \left(\frac{d}{d\vartheta}\right)^{\ell'+m'} (1-\vartheta^2)^{\ell'}. \qquad (9.354)$$

The $P_{\ell'}^{m'}$ are known as associated Legendre polynomials. They evidently satisfy the orthogonality relations

$$\int_{-1}^{1} P_{\ell''}^{m'}(\vartheta)\, P_{\ell'}^{m'}(\vartheta)\, d\vartheta = \frac{2}{2\ell'+1} \frac{(\ell'+m')!}{(\ell'-m')!} \delta_{\ell''\ell'} \qquad (9.355)$$

Comparison of eqs. (9.346) and (9.353) yields us the following identity :

$$(-1)^{m'} (1-\vartheta^2)^{\frac{m'}{2}} \left(\frac{d}{d\vartheta}\right)^{\ell'+m'} (1-\vartheta^2)^{\ell'} = \frac{(\ell'+m')!}{(\ell'-m')!} (1-\vartheta^2)^{-\frac{m'}{2}} \left(\frac{d}{d\vartheta}\right)^{\ell'-m'} (1-\vartheta^2)^{\ell'} \qquad (9.356)$$

Wave functions for $m' \leq 0$ are obtained in a similar manner. Making use of (9.184), we have

$$|\ell',m'\rangle = \frac{1}{\sqrt{[\ell'(\ell'+1)][\ell'(\ell'+1)-(-1)(-2)]\cdots[\ell'(\ell'+1)-(-|m'|+1)(-|m'|)]}}\, L_-^{|m'|}|\ell',0\rangle$$

$$= \sqrt{\frac{(\ell'-|m'|)!}{(\ell'+|m'|)!}}\, L_-^{|m'|} |\ell', 0\rangle, \tag{9.357}$$

or

$$\langle \theta, \varphi | \ell', m' \rangle = (-1)^{\ell'} \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \sqrt{\frac{(\ell'-|m'|)!}{(\ell'+|m'|)!}}$$
$$\times \left[e^{-i\varphi} \left(-\frac{\partial}{\partial\theta} + i \cot\theta \frac{\partial}{\partial\varphi} \right) \right]^{|m'|} P_{\ell'}(\cos\theta)$$
$$= (-1)^{\ell'+|m'|} \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \frac{\sqrt{(\ell'-|m'|)!}}{(\ell'+|m'|)!} e^{-i|m'|\varphi} \left(\frac{\partial}{\partial\theta} - |m'| \cot\theta \right) \cdots \left(\frac{\partial}{\partial\theta} - \cot\theta \right) \frac{\partial}{\partial\theta} P_{\ell'}(\cos\theta) \tag{9.358}$$

yielding

$$\langle \theta, \varphi | \ell', m' \rangle = (-1)^{\ell'} \frac{\Delta\omega^{1/2}}{\sqrt{2\pi}} \sqrt{\frac{2\ell'+1}{2}} \frac{\sqrt{(\ell'-|m'|)!}}{(\ell'+|m'|)!} e^{im'\varphi} P_{\ell'}^{|m'|}(\theta), \tag{9.359}$$

We shall close this section with the derivation of an important relation involving the Legendre polynomials, which we shall use in a later chapter. First, consider the following integral :

$$\int_{-1}^{1} \varphi^n e^{ia\varphi} d\varphi = \frac{1}{a^{n+1}} \int_{-a}^{a} y^n e^{iy} dy \qquad (y = a\varphi)$$
$$= \frac{1}{ia^{n+1}} \left[y^n e^{iy} \right]_{-a}^{a} - \frac{n}{ia^{n+1}} \int_{-a}^{a} y^{n-1} e^{iy} dy \qquad \left(\begin{array}{l} u = y^n \\ dv = e^{iy} dy \\ du = n y^{n-1} dy \\ v = \frac{1}{i} e^{iy} \end{array} \right)$$
$$= \frac{e^{ia} - (-1)^n e^{-ia}}{ia} - \frac{n}{i^2 a^{n+1}} \left[y^{n-1} e^{iy} \right]_{-a}^{a} + \frac{n(n-1)}{i^2 a^{n+1}} \int_{-a}^{a} y^{n-2} e^{iy} dy$$
$$= \frac{e^{ia} - (-1)^n e^{-ia}}{ia} - n \frac{e^{ia} - (-1)^{n-1} e^{-ia}}{(ia)^2} + \cdots \tag{9.360}$$

If we are interested only in large values of a, then we have the asymptotic form

$$\int_{-1}^{1} \rho^n e^{ia\rho}\, d\rho \sim \frac{e^{ia} - (-1)^n e^{-ia}}{ia} \qquad (9.361)$$

More generally, if $f(\rho)$ is any polynomial in ρ, we may write

$$\int_{-1}^{1} f(\rho)\, e^{ia\rho}\, d\rho \sim \frac{f(1)e^{ia} - f(-1)e^{-ia}}{ia} \qquad (9.362)$$

We shall now investigate the case in which $f(\rho)$ is a Legendre polynomial. Let us make the change of variables

$$\alpha = \rho - 1, \qquad \rho^2 - 1 = (\rho - 1)(\rho + 1) = \alpha(\alpha + 2) \qquad (9.363)$$

Then we may write

$$P_\ell(\rho) = P_\ell(1 + \alpha) = \frac{1}{2^\ell \ell!}\left(\frac{d}{d\alpha}\right)^\ell \alpha^\ell (\alpha + 2)^\ell$$
$$= \frac{1}{2^\ell \ell!}\left(\ldots + \ell! 2^\ell\right) \qquad (9.364)$$

It is clear that

$$P_\ell(1) = \left[P_\ell(1 + \alpha)\right]_{\alpha = 0} = 1. \qquad (9.365)$$

Again, making the change of variables

$$\beta = \rho + 1, \qquad \rho^2 - 1 = (\rho - 1)(\rho + 1) = \beta(\beta - 2), \qquad (9.366)$$

we may write

$$P_\ell(\rho) = P_\ell(1 - \beta) = \frac{1}{2^\ell \ell!}\left(\frac{d}{d\beta}\right)^\ell \beta^\ell (\beta - 2)^\ell$$
$$= \frac{1}{2^\ell \ell!}\left(\ldots + (-1)^\ell \ell! 2^\ell\right), \qquad (9.367)$$

so that

$$P_{\ell'}(-1)=\left[P_{\ell'}(1-\beta)\right]_{\beta=0}=(-1)^{\ell'}. \tag{9.368}$$

Hence finally

$$\int_{-1}^{1}P_{\ell'}(\rho)e^{i\alpha\rho}d\rho\sim\frac{e^{i\alpha}-(-1)^{\ell'}e^{-i\alpha}}{i\alpha}. \tag{9.369}$$

Since the Legendre polynomials are a complete set of functions on the interval -1 to 1, the function $e^{i\alpha\rho}$ may be expanded in the form

$$e^{i\alpha\rho}=\sum_{\ell'=0}^{\infty}A_{\ell'}(\alpha)P_{\ell'}(\rho). \tag{9.370}$$

The coefficients $A_{\ell'}(\alpha)$ may be determined by means of the orthogonality relations (9.350),

$$A_{\ell'}(\alpha)=\frac{2\ell'+1}{2}\int_{-1}^{1}P_{\ell'}(\rho)e^{i\alpha\rho}d\rho\sim\frac{2\ell'+1}{2}\frac{e^{i\alpha}-(-1)^{\ell'}e^{-i\alpha}}{i\alpha}. \tag{9.371}$$

We therefore have the asymptotic expansion

$$e^{i\alpha\rho}=\sum_{\ell'=0}^{\infty}\frac{2\ell'+1}{2i\alpha}\left[e^{i\alpha}-(-1)^{\ell'}e^{-i\alpha}\right]P_{\ell'}(\rho). \tag{9.372}$$

a) Systems with bound stationary states.

Consider a system which is described by a time independent Hamiltonian function H. H is a constant of motion. Let us work in the Heisenberg representation, and let us introduce a complete set of orthonormal Heisenberg state vectors $|\alpha'\rangle$. Suppose the operator H is numbered among the α' so that the $|\alpha'\rangle$ are eigenvectors of H. As we have seen, the eigenvalues H' of H may range either over a discrete set, a continuous range, or both.

Let F be an arbitrary dynamical observable. Let $\dot{F}$ denote its time derivative. Let us consider the expectation value

of $\dot{F}$ in one of the energy eigenstates $|\alpha\rangle$. We have

$$\langle \alpha' | \dot{F} | \alpha \rangle = \frac{i}{\hbar} \langle \alpha' | H F - F H | \alpha \rangle = \frac{i}{\hbar} (H' - H) \langle \alpha' | F | \alpha \rangle = 0 \qquad (9.374)$$

This says that the expectation value of F in the energy eigenstate vanishes. But this is clearly nonsense. For, suppose the system consists of a single particle of definite energy being scattered by a center of force, and suppose F is the particle position variable x. Then $\dot{x} = \frac{1}{m} p$, where m is the mass and p the momentum of the particle, and eq. (9.373) says that $\langle p \rangle = 0$, but we know that the particle starts initially with a definite energy, equal to E, and a corresponding definite momentum p'. We know that the expectation value of p is overwhelmingly likely to be approximately p', and not zero.

We are here faced with a contradiction analogous to that considered in the discussion of the Zeno particle (see (9.24)). We shall avoid the contradiction by the same trick as used previously, namely, we shall regard H as the limiting form of a complex operator as its imaginary part tends to zero. That is, we shall still have

$$H |\alpha'\rangle = H' |\alpha'\rangle, \qquad (9.374)$$

where H' is real, but we shall write

$$H^* = H + i\epsilon, \qquad (9.375)$$

where ϵ is a small real quantity which tends to zero in the limit after all algebraic work involving H has been completed. We may then write (9.373) in the corrected form

$$\langle \alpha'|\dot{F}|\alpha\rangle = -\frac{1}{i\hbar}\langle \alpha'|HF - FH|\alpha\rangle$$
$$= -\frac{1}{i\hbar}\langle \alpha'|H^*F - i\varepsilon F - FH|\alpha\rangle$$
$$= -\frac{1}{i\hbar}(H' - i\varepsilon - H')\langle \alpha'|F|\alpha\rangle$$
$$= \frac{1}{\hbar}\langle \alpha'|\varepsilon F|\alpha\rangle \qquad (9.376)$$

Although we shall here not go through the analysis of an actual case*, we can say that when H' comes from a continuum of values, $|\alpha'\rangle$ will be such that the expression on the right of (9.376) does not necessarily vanish in the limit $\varepsilon \to 0$.

When H' is one of a discrete set of eigenvalues, however, simple matrix arithmetic applies, and eq. (9.375) holds. The system is then, moreover, when viewed classically, an ordinary multiply periodic system. The wave function of the system is then, in general, quadratically integrable; its non-vanishing values being more or less limited to a bounded region of configuration space, and the system itself is said to be in a bound stationary state. Thus, when a system is in a bound stationary state the time derivative of the expectation value of any observable is zero. This corresponds classically to the fact that the multiple Fourier series for the time derivative of any dynamical observable contains no constant term.

The constancy of mean values was originally responsible for the introduction of the term "stationary state". Nowadays, however, the expression "stationary state" is used to denote any eigenstate of the Hamiltonian function. We must therefore distinguish between "bound" and "unbound" stationary states.

---

* The use of small imaginary parts in operators appearing in situations similar to the above will be more fully discussed in the chapters on perturbation theory and scattering.

We shall now derive an important theorem for bound stationary states. Consider a system which has a Hamiltonian function of the form

$$H = \sum_i \frac{1}{2m_i} p_i^2 + V \tag{9.377}$$

where V is a function of the ρ^i. We have

$$\dot{\rho}^i = \frac{1}{i\hbar}[\rho^i, H] = \frac{1}{m_i} p_i, \qquad \text{N.S.,} \tag{9.378}$$

$$\dot{p}_i = \frac{1}{i\hbar}[p_i, H] = -\frac{\partial V}{\partial \rho^i}. \tag{9.379}$$

Hence

$$\begin{aligned} \frac{d}{dt}\sum_i p_i \rho^i &= \sum_i (\dot{p}_i \rho^i + p_i \dot{\rho}^i) \\ &= \sum_i \left(\frac{1}{m_i} p_i^2 - \rho_i \frac{\partial V}{\partial \rho^i}\right) \\ &= \sum 2H - 2V - \sum_i \rho^i \frac{\partial V}{\partial \rho^i}. \end{aligned} \tag{9.380}$$

Taking the expectation value of this equation in a bound stationary state, we obtain

$$\begin{aligned} 0 &= \sum_i \frac{1}{m_i} \langle \alpha' | p_i^2 | \alpha' \rangle - \sum_i \langle \alpha' | \rho^i \frac{\partial V}{\partial \rho^i} | \alpha' \rangle \\ &= 2H' - \langle \alpha' | 2V + \sum_i \rho^i \frac{\partial V}{\partial \rho^i} | \alpha' \rangle \end{aligned} \tag{9.381}$$

This is known as the virial theorem.

f) f) Kepler motion.

We may here use many of the results of section d). In the present case equation (9.395) becomes (see (1.252))

$$H = \frac{1}{2m} p^2 - \frac{\kappa e^2}{r}. \tag{9.382}$$

We have the following dynamical equations

$$L_i = \frac{1}{i\hbar}[L_i, H] = 0 \tag{9.383}$$

$$\dot{p}_i = \frac{1}{i\hbar}[p_i, H] = -\kappa e^2 \frac{x_i}{r^3},$$

$$\frac{d}{dt}\frac{x_i}{r} = \frac{1}{i\hbar}\left[\frac{x_i}{r}, H\right] = \frac{1}{2im\hbar}\left[\frac{x_i}{r}, p^2\right]$$

$$= \frac{1}{2m}\left\{\frac{\delta_{ik}}{r} - \frac{x_i x_k}{r^3}, p_k\right\}$$

$$= -\frac{1}{2m}(\delta_{ik}\delta_{jm} - \delta_{im}\delta_{jk})\left\{\frac{x_j x_k}{r^3}, p_m\right\}$$

$$= -\frac{1}{2m}\varepsilon_{ijk}\varepsilon_{klm}\left\{\frac{x_j}{r^3}, x_k p_m\right\}$$

$$= -\frac{1}{2m}\varepsilon_{ijk}\left\{\frac{x_j}{r^3}, L_k\right\}$$

$$= -\frac{1}{2m r^3}\varepsilon_{ijk}\{x_j, L_k\}. \tag{9.384}$$

If we now introduce the operators

$$C_i = \frac{1}{m\kappa e^2} B_i + \frac{x_i}{r}, \tag{9.385}$$

where the $\mathcal{B}_i$ are given by (9.309); we find

$$\dot{\mathcal{E}}_i = -\frac{1}{2mZe^2}\varepsilon_{ijk}\{\dot{p}_j, L_k\} + \frac{d}{dt}\frac{x_i}{r}$$
$$= \frac{1}{2m}\varepsilon_{ijk}\left\{\frac{x_j}{r^3}, L_k\right\} + \frac{d}{dt}\frac{x_i}{r} = 0. \qquad (9.386)$$

That is, the vector $\mathcal{E}$ is a constant of the motion.

Multiplying equation (9.385) on the left by x_i and making use of (9.310), we get

$$x_i\mathcal{E}_i = -\frac{1}{mZe^2}\varepsilon_{ijk}x_i p_j L_k + \frac{i\hbar}{mZe^2}x_i p_i + \hbar$$
$$= -\frac{1}{mZe^2}L^2 + \frac{i\hbar}{mZe^2}x_i p_i + \hbar. \qquad (9.387)$$

Taking the Hermitian adjoint, we have

$$\mathcal{E}_i x_i = -\frac{1}{mZe^2}L^2 - \frac{i\hbar}{mZe^2}p_i x_i + \hbar, \qquad (9.388)$$

and hence

$$\frac{1}{2}\{x_i\mathcal{E}_i\} = -\frac{1}{mZe^2}L^2 + \frac{i\hbar}{2mZe^2}[x_i, p_i] + \hbar$$
$$= -\frac{1}{mZe^2}\left(L^2 + \frac{3}{2}\hbar^2\right) + \hbar. \qquad (9.389)$$

In the classical limit $\hbar \to 0$ this equation takes the form

$$r\cdot\mathcal{E} = -\frac{L^2}{mZe^2} + r, \qquad (9.390)$$

or

$$r\left(1 - |\mathcal{E}|\cos\theta\right) = \frac{L^2}{mZe^2}, \qquad (9.391)$$

where θ is the angle between r and $\mathcal{E}$. Equation (9.391) is the equation for the orbit of the particle. $|e|$ is seen to be the eccentricity of the orbit, and $\mathcal{E}$ is proportional to the vector which points from the center of force at the focus to the center of the orbit.

We shall now make use of the fact that $\mathcal{E}$ is a constant of the motion to construct a time independent algebra involving it and other constants of the motion. In this way we shall be led to the eigenvalues of the energy operator. From eqs. (9.303, 304), (9.317, 318) and (9.385) we get

$$\mathcal{E}_i L_i = L_i \mathcal{E}_i = 0 \tag{9.392}$$

With the help of (9.312) and (9.314) we get

$$\begin{aligned} \mathcal{E}^2 &= \frac{1}{m^2 Z^2 e^4} B^2 + \frac{1}{mZc^2}\left\{B_i, \frac{x_i}{r}\right\} p_i + 1 \\ &= \frac{L^2 + \hbar^2}{m^2 Z^2 e^4} p^2 - \frac{2}{mZc^2 r}(L^2 + \hbar^2) + 1 \\ &= \frac{2}{mZ^2 c^4}(L^2 + \hbar^2) H + 1. \end{aligned} \tag{9.393}$$

Finally, with the help of (9.316) and (9.319), we get

$$\begin{aligned} [\mathcal{E}_i, \mathcal{E}_j] &= \frac{1}{m^2 Z^2 e^4}[B_i, B_j] + \frac{1}{mZc^2}\left([B_i, \frac{x_j}{r}] - [B_j, \frac{x_i}{r}]\right) \\ &= \frac{-i\hbar}{m^2 Z^2 e^4} \epsilon_{ijk} L_k p^2 + \frac{2i\hbar}{mZc^2 r} \epsilon_{ijk} L_k \\ &= \frac{-2i\hbar}{mZ^2 c^4} \epsilon_{ijk} L_k H. \end{aligned} \tag{9.394}$$

We shall be interested in only those matrix elements of the above operators which corresponds to bound stationary states. Under these circumstances the virial theorem (9.381) applies. For the present case we have

$$V = -\frac{Ze^2}{r} \tag{9.395}$$

and

$$x_i \frac{\partial V}{\partial x_i} = Ze^2 x_i \frac{x_i}{r^3} = -V, \tag{9.396}$$

so that

$$2H' = \langle \alpha' | V | \alpha' \rangle = -\langle \alpha' | x_i \frac{\partial V}{\partial x_i} | \alpha' \rangle$$
$$= -\frac{1}{m} \langle \alpha' | p^2 | \alpha' \rangle \tag{9.397}$$

Since p^2 is a positive definite operator we see that the energy eigenvalues corresponding to bound stationary states are all negative.

We introduce the operators

$$I^+ = \frac{1}{2k} L + \frac{1}{2} \sqrt{\frac{mZ^2e^4}{-2k^2H}}\, \epsilon, \tag{9.398}$$

$$I^- = \frac{1}{2k} L - \frac{1}{2} \sqrt{\frac{mZ^2e^4}{-2k^2H}}\, \epsilon. \tag{9.399}$$

These operators satisfy the following relations

$$[I_i^+, I_j^+] = \frac{1}{4k^2} [L_i, L_j] + \frac{1}{4k} \sqrt{\frac{mZ^2e^4}{-2k^2H}} [L_i, \epsilon_j]$$
$$+ \frac{1}{4k} \sqrt{\frac{mZ^2e^4}{-2k^2H}} [\epsilon_i, L_j] - \frac{mZ^2e^4}{8k^2H} [\epsilon_i, \epsilon_j]$$
$$= \frac{i}{4k} \epsilon_{ijk} L_k + \frac{i}{2} \sqrt{\frac{mZ^2e^4}{-2k^2H}}\, \epsilon_{ijk} \epsilon_k + \frac{i}{4k^2} \epsilon_{ijk} L_k$$
$$= i \epsilon_{ijk} I_k^+, \tag{9.400}$$

and similarly

$$[I_i^-, I_j^-] = i \epsilon_{ijk} I_k^- \tag{9.401}$$

and

$$[I_z^+, I_z^-] = 0 \tag{9.402}$$

The operators I_z^+ and I_z^- are seen to satisfy commutation relations analogous to those satisfied by the components of the angular momentum vector. We may therefore treat I_z^+ and I_z^- by the methods of section c) and conclude that the operator I, where

$$\begin{aligned} I(I+1) &= (I^+)^2 = (I^-)^2 \\ &= \frac{1}{4\hbar^2} L^2 - \frac{m\mathcal{E}^2 e^4}{8\hbar^2 H} \\ &= \frac{1}{4\hbar^2} L^2 - \frac{1}{4\hbar^2}(L^2 + \hbar^2) - \frac{m\mathcal{E}^2 e^4}{8\hbar^2 H} \\ &= -\frac{1}{4}\left(1 + \frac{m\mathcal{E}^2 e^4}{2\hbar^2 H}\right), \end{aligned} \tag{9.403}$$

has the eigenvalues

$$I' = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \tag{9.404}$$

Writing

$$I' = \frac{n'-1}{2} \tag{9.405}$$

where $n' = 1, 2, 3, \ldots,$ (9.406)

we have

$$\begin{aligned} -\frac{1}{4}\left(1 + \frac{m\mathcal{E}^2 e^4}{2\hbar^2 H'}\right) &= I'(I'+1) \\ &= \frac{n'-1}{2}\,\frac{n'+1}{2} = -\frac{1}{4}(1 - n'^2), \end{aligned} \tag{9.407}$$

from which we obtain at once the energy eigenvalues (cf. (5.30)):

$$H' = -\frac{m Z^2 e^4}{2\hbar^2} \frac{1}{n'^2} \tag{9.408}$$

Since H' is negative and $\hbar^2$ is a positive definite operator, we see from (9.403) that we must have

$$I'(I'+1) \geq \frac{1}{4}\ell'(\ell'+1) \tag{9.409}$$

or

$$n'^2 - 1 \geq \ell'(\ell'+1) \tag{9.410}$$

which gives us the restriction (cf. (5.28))

$$\ell' \leq n' - 1 . \tag{9.411}$$

Since the operators I, I_3^+, I_3^- can be taken as a complete set of commuting real operators for the system, it is evident that the degeneracy of the energy level H' is

$$(2I' + 1)^2 = n'^2 \qquad \text{(cf. (5.34)}$$

Bibliography

P.A.M. Dirac, The Principles of Quantum Mechanics, Oxford, (1947)

M. Born und P. Jordan, Elementare Quantenmechanik, Julius Springer, Berlin, (1930).

D. Bohm, Quantum Theory, New-York, Prentice-Hall, Inc., (1951)

COURS DE L'ECOLE D'ETE DE PHYSIQUE THEORIQUE

QUANTUM MECHANICS

Bryce S. DeWITT

Radiation Laboratory

University of California

★ ★ ★

★ ★

LES HOUCHES

ETE 1953

Q U A N T U M M E C H A N I C S

Lecture Notes by
Bryce S. DE WITT
Radiation Laboratory
University of California

Chapter X.

Université de Grenoble
Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

10.- QUANTUM PERTURBATION THEORY.

As in the classical perturbation theory of chapter 2 we consider a system for which the Hamiltonian function has the form

$$H = H_0 + H_1 \tag{10.1}$$

where the function H_0 is of a sufficiently simple form that the complete general solution of the equations of motion of the system for which it is the Hamiltonian function is known.

Let us start from the Schrödinger representation in which the state vector $|t\rangle$ satisfies the differential equation (see (8.43))

$$i\hbar \frac{d}{dt}|t\rangle = H_S(t)|t\rangle = [H_{0S}(t) + H_{1S}(t)]|t\rangle, \tag{10.2}$$

where the subscript S is used to denote operators in the Schrödinger representation, and the operators H_{0S} and H_{1S} are written $H_{0S}(t)$, $H_{1S}(t)$ to indicate any possible explicit time dependence which they may have. Let us then introduce an operator $U_0(t'',t')$ which satisfies the integral equation

$$U_0(t'',t') = 1 + \frac{1}{i\hbar}\int_{t'}^{t''} H_{0S}(t)\,U_0(t,t')\,dt \tag{10.3}$$

From the discussion of the similar operator $U(t'',t')$ on pages VII-6 to VIII-10 it is evident that $U_0(t'', t')$ is a unitary operator which satisifes the relations

$$U_0(t'',t')\,U_0(t',t''') = U_0(t'',t'''), \tag{10.4}$$

$$U_0^*(t'',t') = U_0(t',t''), \tag{10.5}$$

the differential equations

$$i\hbar \frac{\partial}{\partial t''} U_0(t'',t') = H_{0S}(t'')\, U_0(t'',t') \tag{10.6}$$

$$-i\hbar \frac{\partial}{\partial t'} U_0(t'',t') = U_0(t'',t')\, H_{0S}(t'), \tag{10.7}$$

and the boundary condition

$$U_0(t',t') = 1. \tag{10.8}$$

Let us now introduce the following state vector

$$|t\rangle_I = U_0(t_0,t)|t\rangle, \quad |t\rangle = U_0(t,t_0)|t\rangle_I \tag{10.9}$$

where t_0 is the fixed point of time used in passing from the Heisenberg to the Schrödinger representation. The expectation value of any observable F, which is not explicitly dependent on the time, may be expressed in terms of this state vector as follows :

$$\bar{F} = \langle t|F|t\rangle = \langle t|F_S|t\rangle = {}_I\langle t|U_0(t_0,t)F_S U_0(t,t_0)|t\rangle_I$$
$$= {}_I\langle t|F_I(t)|t\rangle_I \tag{10.10}$$

where

$$F_I(t) = U_0(t_0,t)F_S U(t,t_0) \tag{10.11}$$

Although the Schrödinger operator F_S is time-independent, the operator F_I possesses a time variation. Using eqs. (10,6, 7), we have

$$\dot{F}_I = \dot{U}_0(t_0,t)F_S U(t,t_0) + U(t_0,t)F_S \dot{U}(t,t_0)$$
$$= \frac{1}{i\hbar} U_0(t_0,t)[-H_{0S}(t)F_S + F_S H_{0S}(t)]U(t,t_0)$$
$$= \frac{1}{i\hbar}[F_I, H_{0I}] \qquad (10.12)$$

where

$$H_{0I}(t) = U_0(t_0,t)\, H_{0S}(t)\, U_0(t,t_0). \qquad (10.13)$$

Equations (10.12) are simply the general equations of motion of a system possessing the Hamiltonian function H_0, and their solutions will, as we have remarked, be well known. Therefore the t time behavior of the operators F_I will henceforth be regarded as given. The problem then remains of determining the time behavior of the state vectors $|t\rangle_I$. Using equations (8.43), (10.7) and (10.9) we have

$$i\hbar \frac{d}{dt}|t\rangle_I = i\hbar\left[\dot{U}_0(t_0,t)|t\rangle + U_0(t_0,t)\frac{d}{dt}|t\rangle\right]$$
$$= U_0(t_0,t)\left[-H_{0S}(t) + H_S(t)\right]|t\rangle$$
$$= U_0(t_0,t)\, H_{1S}(t)\, U_0(t,t_0)|t\rangle_I = H_{1I}(t)|t\rangle_I. \qquad (10.14)$$

The description of a physical system in terms of the state-vectors $|t\rangle_I$ and the operators F_I is called the interaction representation. The function H_1 is usually known as the interaction Hamiltonian, and the operator H_{1I} is called the interaction operator.

It is evident that the Heisenberg, Schrödinger and interaction representations all coincide at the time t_0 :

$$|t_0\rangle_I = |t_0\rangle = |\,\rangle, \qquad (10.15)$$

$$F_I(t_0) = F_S(t_0) = F(t_0). \tag{10.16}$$

Making use of the operator $U_S(t'', t')$ defined on page VIII 10, we may write

$$\begin{aligned} |t''\rangle_I &= U_0(t_0, t'')|t''\rangle = U_0(t_0, t'')U_S(t'', t')|t'\rangle \\ &= U_0(t_0, t'')U_S(t'', t')U_0(t', t_0)|t'\rangle_I \\ &= U_I(t'', t')|t'\rangle_I \end{aligned} \tag{10.17}$$

where

$$U_I(t'', t') = U_0(t_0, t'')U_S(t'', t')U_0(t', t_0). \tag{10.18}$$

The operators $U_I(t'', t')$ satisfy the following relation :

$$\begin{aligned} &U_I(t''', t'')U_I(t'', t') \\ &= U_0(t_0, t''')U_S(t''', t'')U_0(t'', t_0)U_0(t_0, t')U_S(t'', t')U_0(t', t_0) \\ &= U_0(t_0, t''')U_S(t''', t'')U_S(t'', t')U_0(t', t_0) \\ &= U_0(t_0, t''')U_S(t''', t')U_0(t', t_0) = U_I(t''', t'). \end{aligned} \tag{10.19}$$

They also evidently satisfy the differential equations

$$i\hbar \frac{\partial}{\partial t''} U_I(t'', t') = H_{1I}(t'')U_I(t'', t'), \tag{10.20}$$

$$-i\hbar \frac{\partial}{\partial t'} U_I(t'', t') = U_I(t'', t')H_{1I}(t'), \tag{10.21}$$

and the boundary condition

$$U_I(t', t') = 1 \tag{10.22}$$

Equations (10.20-22) may be replaced by the integral e equation

$$U_I(t'',t') = 1 + \frac{1}{ih}\int_{t'}^{t''} H_{1I}(t)\, U_I(t,t')\, dt. \quad (10.23)$$

A formal solution of this equation may be obtained by a process of iteration. One substitutes in the integrand the expression for $U_I(t, t')$ given by the right hand side of (10.23) with t" replaced by t, and so on repeatedly. One obtains

$$U_I(t'',t') = 1 + \frac{1}{ih}\int_{t'}^{t''} H_{1I}(t)\,dt + \left(\frac{1}{ih}\right)^2 \int_{t'}^{t''} dt_2 \int_{t'}^{t''} dt_1 \quad (10.24)$$
$$(H_{1I}(t_2) H_{1I}(t_1) + \dots$$

One may write more compactly

$$U_I(t'',t') = \sum_{n=0}^{\infty} \frac{2}{(2ih)^n} \int_{t'}^{t''} dt_n \int_{t'}^{t''} dt_1 \left(1 + \frac{t_n - t_{n-1}}{|t_n - t_{n-1}|}\right) \dots \left(1 + \frac{t_2 - t_1}{|t_2 - t_1|}\right)$$
$$\times H_{1I}(t_n) \dots H_{1I}(t_1)$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!(ih)^n} \int_{t'}^{t''} dt_1 \int_{t'}^{t''} dt_n\, P[H_{1I}(t_1) \dots H_{1I}(t_n)]$$
$$= P\left[\exp \frac{1}{ih}\int_{t'}^{t''} H_{1I}(t)\,dt\right], \quad (10.25)$$

where P is an operator, known as the time-ordering operator, which arranges the operators standing in the bracket following it in order, left to right, of decreasing values of the times t_i at which the operators are evaluated.

The formal expansion (10.25) may be expected to converge only under restricted conditions. We shall investigate the problem of convergence at a later point. The actual evaluation of the terms of the series (10.25) is in general a difficult matter. However, it frequently happens that the first two or three terms can be evaluated without excessive difficulty and that the sum of these

terms alone is a fair approximation to the operator function $U_I(t'', t')$. Using the resulting expression for $U_I(t'', t')$, one then has a complete solution,, to the degree of approximation in question of the quantum dynamical problem posed by the Hamiltonian function (10.1). The approximation will be good, of course, only if H_1 is small compared to H_0. Indeed, the spirit of perturbation theory rests in the assumption that H_1 <u>is</u> small compared to H_0.

Let us now consider the case in which the function H_0 has no explicit dependence on the time. Then the operator $U_0(t'',t')$ takes the form (cf. (8.46))

$$U_0(t'',t') = e^{\frac{1}{i\hbar} H_{0s}(t''-t')} \qquad (10.26)$$

the operator H_{0s} being a constant independent of the time. Furthermore, from eqs. (1.98-102) we know that the system which has H_0 as a Hamiltonian function possesses n independent constants or integrals of motion which have vanishing Poisson brackets with one another. (n is the number of degrees of freedom of the system). These constants of the motion may be taken to be the p'''_i of the canonical transformation equations (1.101, 102) :

$$p'''_i = p'''_i(p, q), \quad i = 1 \ldots n. \qquad (10.27)$$

Since the function W is time independent we know that the functional expressions (10.27) for the p'''_i, which can be obtained by solving eqs. (1.101, 102), have no explicit time dependence. Moreover, since the p'''_i are canonical momenta, we known that these functions have vanishing Poisson bracket with one another.

In the quantum theory the p'''_i become operators which commute with one another. They constitute, moreover, a complete set of independent commuting operators, since any other independent

function of the ρ's and p's must involve the ρ_i''' and hence have non-vanishing commutator bracket with the p_i'''. We may therefore choose as a basis for a representation of state vectors the simultaneous eigenvectors of the p_i'''. The usefulness of such a basis lies in the fact that the p_i''' are constants of the motion, and hence the basis has physical significance even when we are not working in the Schrödinger representation.

Instead of denoting the members of such a set of constant operators by the letter p, we shall use the letter α. Thus, we shall speak of a complete set of commuting constant operators α_{oi}, $i = 1, \ldots, n$, the subscript o denoting that the operators refer to the system whose Hamiltonian function is H_o. The Hamiltonian function H_o may itself be included among the α_{oi}. It will at any rate be a function of the α_{oi} and will commute with the α_{oi}.

If we now start from the Schrödinger representation for the system whose Hamiltonian function is $H = H_o + H_1$, and pass to the interaction representation, we have using (10.11), (10 (10.26), and the commutativity of H_o with the α_{oi},

$$\alpha_{oiI} = U_o(t_o,t)\,\alpha_{oiS}\,U_o(t,t_o)$$
$$= e^{\frac{1}{i\hbar}H_{oS}(t_o-t)}\,\alpha_{oiS}\,e^{\frac{1}{i\hbar}H_{oS}(t-t_o)} \qquad (10.28)$$
$$= \alpha_{oiS}.$$

That is, the interaction and Schrödinger operators corresponding to the α_{oi} are identical, and without ambiguity we may drop the subscripts I, and S from the expressions α_{oiI}, α_{oiS} and also from H_{oI} and H_{oS}. The set of operators α_{oi} will then be denoted by the collective symbol α_o, and we may write

$$\alpha_{oi}\,|\alpha_o'\rangle = \alpha_{oi}'\,|\alpha_o'\rangle, \qquad (10.29)$$

and, in particular

$$H_0|\alpha_0'\rangle = H_0'|\alpha_0'\rangle, \tag{10.30}$$

or, more generally,

$$f(H_0)|\alpha_0'\rangle = f(H_0')|\alpha_0'\rangle, \tag{10.31}$$

where f is an arbitrary function.

Suppose now that at an initial time t', the system (10.1) is in the Schrödinger state $|\alpha_0'\rangle$. That is

$$|t'\rangle = |\alpha_0'\rangle \tag{10.32}$$

At a subsequent time t" the system will be in a state $|t''\rangle$ given by

$$|t''\rangle = U_S(t'',t')|t'\rangle = U_S(t'',t')|\alpha_0'\rangle. \tag{10.33}$$

We may ask what is the probability that the system will be found, on measurement, to be in the state $|\alpha_0''\rangle$ at the time t". This probability is given by

$$\begin{aligned} P(\alpha_0'',t''|\alpha_0',t') &= |\langle\alpha_0''|t''\rangle|^2 \\ &= |\langle\alpha_0''|U_S(t'',t')|\alpha_0'\rangle|^2 \end{aligned} \tag{10.34}$$

But using eqs. (10.18) and (10.26), we may write

$$\begin{aligned} \langle\alpha_0''|U_S(t'',t')|\alpha_0'\rangle &= \langle\alpha_0''|U_0(t'',t_0)U_I(t'',t_0)U_0(t_0,t')|\alpha_0'\rangle \\ &= \langle\alpha_0''|e^{\frac{1}{i\hbar}H_0(t''-t_0)}U_I(t'',t')e^{\frac{1}{i\hbar}H_0(t_0-t')}|\alpha_0'\rangle \\ &= e^{\frac{1}{i\hbar}[H_0''(t''-t_0)-H_0'(t'-t_0)]}\langle\alpha_0''|U_I(t'',t')|\alpha_0'\rangle, \end{aligned} \tag{10.35}$$

and hence

$$P(\alpha_0'', t''/\alpha_0', t') = |\langle \alpha_0''| U_I(t'', t') |\alpha_0'\rangle|^2 \qquad (10.36)$$

The function $P(\alpha_0'', t''/\alpha_0', t')$ gives the probability that the system which is in the state $|\alpha_0'\rangle$ at the time t' will make a transition to the state $|\alpha_0''\rangle$ by the time t". Equations (10.34) and (10.36) show that this transition probability can be calculated equally well by means of the operator $U_S(t'', t')$ or the operator $U_I(t'', t')$. The series expansion (10.24, 25) may therefore be used to calculate transition probabilities to any desired accuracy. In particular, if $\alpha_0'' \neq \alpha_0'$, we have, to the lowest order of accuracy

$$P(\alpha_0'', t''/\alpha_0', t') \approx \frac{1}{\hbar^2} \left| \int_{t'}^{t''} \langle \alpha_0''| H_{1I}(t) |\alpha_0'\rangle dt \right|^2 \qquad (10.37)$$

Let us now make a further specialization to the case in which the function H_1, like H_0, also has no explicit dependence on the time. The total Hamiltonian function H is then time independent and constant, and we shall have the quantum analog of the time independent classical perturbation theory of chapter 2. This case is of great importance, and we shall now consider it in some detail.

Just as we did for the system whose Hamiltonian function was H_0, we can introduce a complete set of commuting operators which are constants of the motion for the system whose Hamiltonian function is H. H will be a function of the α_i and will commute with them. Therefore their Schrödinger and Heisenberg forms will be identical

$$\alpha_{iS} = U(t, t_0)\, \alpha_i\, U(t_0, t)$$
$$= e^{\frac{1}{i\hbar} H(t-t_0)}\, \alpha_i\, e^{\frac{1}{i\hbar} H(t_0-t)} = \alpha_i, \qquad (10.38)$$

using (8.40) and (8.46). Equation (8.47) is a special case of (10.38). We shall drop the subscript S from the expressions α_{iS} and H_S,

and write, in both the Heisenberg and Schrödinger representations,

$$\alpha_i |\alpha'\rangle = \alpha_i' |\alpha'\rangle. \tag{10.39}$$

In particular

$$H |\alpha'\rangle = H' |\alpha'\rangle, \tag{10.40}$$

or, more generally,

$$f(H) |\alpha'\rangle = f(H') |\alpha'\rangle, \tag{10.41}$$

where f is an arbitrary function.

Suppose now that the system is in an eigenstate of the energy operator H, i.e. in a stationary state (see p. IX.79ff). More particularly, suppose that the Heisenberg state vector of the system is an eigenvector of the α_i. That is,

$$|\rangle = |\alpha'\rangle. \tag{10.42}$$

The Schrödinger state vector then has a very simple form. Using (8.38) and (8.46), we have

$$|t\rangle = U(t, t_0)|\rangle = e^{-\frac{i}{\hbar}H(t-t_0)}|\alpha'\rangle$$
$$= e^{-\frac{i}{\hbar}H'(t-t_0)}|\alpha'\rangle. \tag{10.43}$$

The Schrödinger state vector is seen to remain parallel to the Heisenberg state vector, only differing from it by a phase factor which is periodic with the angular frequency $H'/\hbar$.

Part of the task of time-independent quantum perturbation theory will be to express the eigenvectors and eigenvalues of the operator H in terms (at least approximately) of the known eigenvectors and eigenvalues of the operator H_0. In doing this it turns

out that it will be useful to consider the operator

$$U_I(t,-\infty) = \lim_{t'\to-\infty} U_I(t,t')$$
$$= \lim_{t'\to-\infty} U_0(t_0,t)U_S(t,t')U_0(t',t_0)$$
$$= \lim_{t'\to-\infty} e^{\frac{1}{i\hbar}H_0(t_0-t)} e^{\frac{1}{i\hbar}H(t-t')} e^{\frac{1}{i\hbar}H_0(t'-t_0)} \quad (10.44)$$

in which the operators H and H_0 in the last line are to be understood as being in their constant Schrödinger forms. In passing to the last line of (10.44) we have made use of (10.26) and the result opposite on page VIII.10.

Now it is quite evident that the expression (10.44) is ambiguous, as it is not at all clear what interpretation should be given to the oscillating factors in the limit $t' \to -\infty$. In order to define the operator $U_I(t,-\infty)$ in an unambiguous manner we shall adopt an adiabatic procedure. First observe that from (10.11) and (10.26), we have

$$H_{1I}(t) = U_0(t_0,t)H_{1S}U_0(t,t_0) = e^{\frac{1}{i\hbar}H_0(t_0-t)} H_{1S} e^{\frac{1}{i\hbar}H_0(t-t_0)} \quad (10.45)$$

where H_{1S} is the constant Schrödinger form of the interaction operator. We shall now replace (10.45) by the expression

$$H_{1I}(t) = \begin{cases} e^{\frac{1}{i\hbar}H_0(t_0-t)} H_{1S} e^{\frac{\varepsilon}{\hbar}(t-t_0)} e^{\frac{1}{i\hbar}H_0(t-t_0)}, & t \leq t_0 \\ e^{\frac{1}{i\hbar}H_0(t_0-t)} H_{1S} e^{\frac{1}{i\hbar}H_0(t-t_0)}, & t \geq t_0 \end{cases} \quad (10.46$$

and impose an implicit limiting convention whereby the limit $\varepsilon \to +0$ is to be taken after all analytical work involving the operator $H_{1I}(t)$ has been carried out. Expression (10.46) is just what $H_{1I}(t)$ would indeed be if the perturbation function H_1 were switched on adiabatically by means of a very small exponential dam-

ping factor and built up to its final value at the time t_0.

Let us try to calculate the operator $U_I(t, -\infty)$ by using (10.46) and the formal expansion (10.25). Introducing the integration variables $t'_1 \ldots t'_n$, given by

$$\left.\begin{aligned} t'_1 &= t_1 - t_2 & t_1 &= t'_1 + t'_2 + \ldots + t'_n + t_0 \\ t'_2 &= t_2 - t_3 & t_2 &= t'_2 + \ldots + t'_n + t_0 \\ &\cdots\cdots & &\cdots\cdots \\ t'_{n-1} &= t_{n-1} - t_n \\ t'_n &= t_n - t_0 & t_n &= t'_n + t_0 \end{aligned}\right\} \qquad (10.47)$$

we have, for $t \leq t_0$,

$$U_I(t,-\infty) = \sum_{n=0}^{\infty} \frac{2}{(2\pi\hbar)^n} \int_{-\infty}^{t} dt_n \int_{-\infty}^{t_n} dt_{n-1} \cdots \int_{-\infty}^{t_2} dt_1 \left(1 + \frac{t_n - t_{n-1}}{|t_n - t_{n-1}|}\right)$$

$$\cdots \left(1 + \frac{t_2 - t_1}{|t_2 - t_1|}\right) \times e^{\frac{1}{i\hbar} H_0 (t_0 - t_n)} H_{1S}\, e^{\frac{1}{i\hbar}(H_0 + \varepsilon)(t_n - t_{n-1})}$$

$$\times \cdots e^{\frac{1}{i\hbar} H_0 (t_2 - t_1)} H_{1S}\, e^{\frac{1}{i\hbar}(H_0 + \varepsilon)(t_1 - t_0)}$$

$$= \sum_{n=0}^{\infty} \frac{2}{(2\pi\hbar)^n} \int_{-\infty}^{t} dt_n \int_{-\infty}^{t_n} dt_{n-1} \cdots \int_{-\infty}^{t_2} dt_1 \left(1 + \frac{t_n - t_{n-1}}{|t_n - t_{n-1}|}\right)$$

$$\cdots \left(1 + \frac{t_2 - t_1}{|t_2 - t_1|}\right) \times e^{\frac{1}{i\hbar} H_0 (t_0 - t_n)} H_{1S}\, e^{\frac{1}{i\hbar} H_0 (t_n - t_{n-1})}$$

$$H_{1S}\, e^{\frac{1}{i\hbar} H_0 (t_{n-1} - t_{n-2})} \times \cdots \times e^{\frac{1}{i\hbar} H_0 (t_2 - t_1)} H_{1S}$$

$$e^{\frac{1}{i\hbar} H_0 (t_1 - t_0)}\, e^{\varepsilon/\hbar (t_1 + t_2 + \cdots + t_n - n t_0)}$$

$$= \sum_{n=0}^{\infty} \frac{2}{(2i\hbar)^n} \int_{-\infty}^{t-t_0} dt'_n \int_{-\infty}^{\infty} dt'_{n-1} \cdots \int_{-\infty}^{\infty} dt'_1 \left(1 - \frac{t'_{n-1}}{|t'_{n-1}|}\right) \cdots \left(1 - \frac{t'_1}{|t'_1|}\right)$$
$$\times e^{-\frac{1}{i\hbar}(H_0 - in\varepsilon)t'_n} H_{1S}\, e^{-\frac{1}{i\hbar}(H_0 - i(n-1)\varepsilon)t'_{n-1}} H_{1S} \times \cdots$$
$$\cdots \times e^{-\frac{1}{i\hbar}(H_0 - i\varepsilon)t'_1} H_{1S}\, e^{\frac{1}{i\hbar}H_0(t'_1 + t'_2 + \cdots + t'_n)}$$

$$= 1 + \sum_{n=1}^{\infty} \sum_{\alpha'_0} \frac{1}{(i\hbar)^n} \int_{-\infty}^{t-t_0} dt'_n \int_{-\infty}^{t'_n} dt'_{n-1} \cdots dt'_1\, e^{-\frac{1}{i\hbar}(H_0 - H'_0 - in\varepsilon)t}$$
$$H_{1S} \times e^{\frac{1}{i\hbar}(H_0 - H'_0 - i(n-1)\varepsilon)t'_{n-1}} H_{1S} \times \cdots \times e^{\frac{1}{i\hbar}(H_0 - H'_0 - i\varepsilon)t'_1} H_{1S} |\alpha'_0\rangle$$

$$= 1 + \sum_{n=1}^{\infty} \sum_{\alpha'_0} (-1)^n e^{-\frac{1}{i\hbar}(H_0 - H'_0 - in\varepsilon)(t-t_0)}$$
$$\frac{1}{H_0 - H'_0 - in\varepsilon} H_{1S} \times \frac{1}{H_0 - H'_0 - i(n-1)\varepsilon} H_{1S} \times \cdots \times \frac{1}{H_0 - H'_0 - i\varepsilon}$$
$$\times H_{1S} |\alpha'_0\rangle \langle \alpha'_0| \qquad (10.48)$$

The series (10.48) may readily be shown to be absolutely convergent as long as ε remains finite. For, let us write (10.48) in the form

$$U_I(t, -\infty) = 1 + \sum_{\alpha'_0} U(\alpha'_0) \qquad (10.49)$$

where

$$U(\alpha'_0) = \sum_{n=1}^{\infty} U_n(\alpha'_0) \qquad (10.50)$$

with

$$U_n(\alpha_0') = (-1)^n e^{-\frac{1}{i\hbar}(H_0 - H_0' - in\varepsilon)(t-t_0)} \frac{1}{H_0 - H_0' - in\varepsilon} H_{1S} \times \cdots \times \frac{1}{H_0 - H_0' - i\varepsilon} H_{1S} |\alpha_0'\rangle \langle \alpha_0'| \quad (10.51)$$

We evidently have the recurrence formula

$$U_{n+1}(\alpha_0') = R_{n+1}(\alpha_0') U_n(\alpha_0') \quad (10.52)$$

where

$$R_{n+1}(\alpha_0') = -e^{-\frac{1}{i\hbar}(H_0 - H_0' - i(n+1)\varepsilon)(t-t_0)} \frac{1}{H_0 - H_0' - i(n+1)\varepsilon} H_{1S} e^{\frac{1}{i\hbar}(H_0 - H_0' - in\varepsilon)(t-t_0)} \quad (10.53)$$

$R_{n+1}(\alpha_0')$ may be interpreted as the ratio of the (n+1) st term of the series (10.50) to the n th term. As long as ε remains finite, we have

$$\lim_{n \to \infty} R_{n+1}(\alpha_0') = 0 \quad (10.54)$$

Hence the series (10.50) is absolutely convergent. The only remaining question is whether the operator $U_I(t, -\infty)$ defined as plus the sum over all α_0' of the operators $U(\alpha_0')$ is well-defined. That it <u>is</u> well-defined is seen at once by writing

$$U_I(t, -\infty) |\alpha_0'\rangle = |\alpha_0'\rangle + U(\alpha_0') |\alpha_0'\rangle \quad (10.55)$$

It is well-defined since its effect on each member of the complete set of vectors $|\alpha_0'\rangle$ is well defined.

In the passage to the limit $\varepsilon \to +0$ two questions arise. The first is the question of whether or not the well-defined operator $U_I(t, -\infty)$ approaches a well-defined limit. We shall see later that it often does <u>not</u>. We shall, in fact, find that $U_I(t, -\infty)$ remains perfectly well behaved in the limit. The second question then arises,

namely : does the series (10.48) converge in the limit $\varepsilon \to +0$? We shall see later that even though $U_I(t, -\infty)$ itself remains well defined, its series representation (10.48) may no longer converge in the limit. Under these circumstances we shall want to be able to discuss the properties of the operator $U_I(t, -\infty)$ independently of series expansions. In order to do this we must first develop our formalism somewhat further.

Using eq. (2.18) of the Appendix we may write formally, in the limit $\varepsilon \to +0$,

$$U_I(t,-\infty) = 1 + \sum_{n=1}^{\infty} \sum_{\alpha_0} (-2\pi i)^n e^{-\frac{1}{i\hbar}(H_0 - H_0')(t-t_0)} \times \left[\delta(H_0 - H_0')H_{IS}\right]^n |\alpha_0\rangle\langle\alpha_0| \quad (10.55)$$

Expressions (10.55) suggests that we may write

$$U_I(t,-\infty) = e^{-\frac{1}{i\hbar}H_0(t-t_0)} U_I(t_0,-\infty) e^{\frac{1}{i\hbar}H_0(t-t_0)} \quad (10.56)$$

together with the formal expansion

$$U_I(t_0,-\infty) = 1 + \sum_{n=1}^{\infty} \sum_{\alpha_0} (-2\pi i)^n \left[\delta(H-H_0')H_{IS}\right]^n |\alpha_0\rangle\langle\alpha_0| \quad (10.57)$$

Now, one may see directly from (10.48) that, for finite ε, the operator $U_I(t, -\infty)$ satisfies the boundary condition

$$\lim_{t\to-\infty} U_I(t,-\infty) = 1. \quad (10.58)$$

The fulfillment of this condition in the limit $\varepsilon \to +0$ is not immediately apparent from the forms (10.55) and (10.56). We see that under the terms of the limiting procedure which we have introduced here we must make the formal identification

$$\lim_{t\to-\infty} e^{-\frac{1}{i\hbar}H_0(t-t_0)} U_I(t_0,-\infty) e^{\frac{1}{i\hbar}H_0(t-t_0)} = 1. \quad (10.59)$$

Expressions (10.56, 57), which hold in the limit $\epsilon \to +0$, are valid for values of t greater than t_0 as well as for $t \leqq t_0$, since the exponential damping factor no longer appears. They are formal solutions of the integral equation satisfied by $U_I(t, -\infty)$ in the limit, namely

$$U_I(t,-\infty) = 1 + \frac{1}{i\hbar}\int_{-\infty}^{t} H_{1I}(t')\,U_I(t',-\infty)\,dt' \qquad (10.60)$$

Now, relation (10.56) should hold independently of any series expansion. Hence, substituting it in (10.60) and making use of (10.45) we obtain the following integral equation for $U_I(t_0, -\infty)$:

$$U_I(t_0,-\infty) = 1 + \frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{-\frac{1}{i\hbar}H_0(t'-t_0)}\,H_{1S}\,U_I(t_0,-\infty)\,e^{\frac{1}{i\hbar}H_0(t'-t_0)}\,dt'$$

$$= 1 + \frac{1}{i\hbar}\sum_{\alpha_0'}\int_{-\infty}^{0} e^{-\frac{1}{i\hbar}(H_0-H_0')t}\,dt\; H_{1S}\,U_I(t_0,-\infty)\,|\alpha_0'\rangle\langle\alpha_0'|$$

$$= 1 - 2\pi i \sum_{\alpha_0'} \delta_-(H_0-H_0')\,H_{1S}\,U_I(t_0,-\infty)\,|\alpha_0'\rangle\langle\alpha_0'| \qquad (10.61)$$

It is evident that the expansion (10.57) is obtained by iteration of (10.61). Using (10.61) directly, however, we may now avoid discussing $U_I(t, -\infty)$ in terms of series expansions.

Let us introduce the following state vector :

$$|\;\rangle_{\alpha_0'} \equiv U_I(t_0,-\infty)\,|\alpha_0'\rangle. \qquad (10.62)$$

From (10.61) it is evident that this state vector satisfies the following integral equation :

$$|\;\rangle_{\alpha_0} = |\alpha_0'\rangle - 2\pi i\,\delta_-(H_0-H_0')\,H_{1S}\,|\;\rangle_{\alpha_0}$$

$$= |\alpha_0'\rangle - \frac{1}{H_0-H_0'-i\epsilon}\,H_{1S}\,|\;\rangle_{\alpha_0}, \qquad (10.63)$$

where use has been made of eq. (2.18) of the Appendix. Multiplying this equation on the left by $H_o - H_o' - \epsilon$, taking the implied limit $\epsilon \to 0$, and using the fact that $(H_o - H_o')|\alpha_o'\rangle = 0$, we obtain

$$(H_o - H_o')|\rangle_{\alpha_o'} = -H_{1S}|\rangle_{\alpha_o'} \qquad (10.64)$$

or

$$H_S|\rangle_{\alpha_o'} = H_o'|\rangle_{\alpha_o'} \qquad (10.65)$$

That is, the vector $|\rangle_{\alpha_o'}$ is an eigenvector of the <u>total Hamiltonian operator H_S</u> (or H) corresponding to the eigenvalue H_o'. We see, therefore, that if the operator H possesses any eigenvalues in common with the operator H_o, it <u>may</u> be possible to obtain the eigenvectors of H corresponding to those eigenvalues through application of the operator $U_I(t_o, -\infty)$. To determine under just what circumstances it <u>will</u> be possible to obtain eigenvectors of H in this manner, we must take a closer look at the limiting processes contained in the above formalism and investigate the conditions under which it is legitimate to write purely formal equations like (10.61) and (10.63).

We can perhaps best begin by considering a situation in which equations (10.61) and (10.63) are <u>not</u> legitimate, namely that which exists for a system whose energy levels are discrete and which possesses bound stationary states as defined in section c) of chapter 9. The energy levels of the perturbed system will then in general differ from those of the unperturbed system. That is, the perturbation function H_1 will cause a shift in the eigenvalues of H from those of H_o, and equation (10.65) certainly cannot hold in these circumstances if $|\rangle_{\alpha_o'}$ is a finite vector. We are led to expect, therefore, that in the case of bound systems, either one or other of the following situations holds :

$$U_I(t_0, -\infty)|\alpha_0'\rangle = 0 \qquad (10.66)$$

or

$$U_I(t_0, -\infty)|\alpha_0'\rangle = \infty . \qquad (10?67)$$

The second of these equations is merely meant to indicate that $|\ \rangle_{\alpha_0'}$ is some sort of a non-convergent vector. We shall see presently that the actual situation is usually characterized by (10.67) rather than by (10.66). That is, the operator $U_I(t_0, -\infty)$ will be found to be non-convergent in the limit $\varepsilon \to 0$. To find out how we must reformulate the preceding theory so as correctly to include bound systems in the scheme, let us first review the old-fashioned methods of bound-state perturbation theory. Let us consider the case in which the unperturbed system is non-degenerate. Then the labels α_0' reduce to the single lable H_0', and one customarily expands the eigenvectors and eigenvalues of the total Hamiltonian operator in the form

$$|\ \rangle = |H_0'\rangle + |\ \rangle_1 + |\ \rangle_2 + |\ \rangle_3 + \ldots \qquad (10.68)$$

$$H' = H_0' + E_1 + E_2 + E_3 + \ldots \qquad (10.69)$$

where the successive terms in the expansions refer to quantities of increasing order in a constant to which the perturbation function H_1 is imagined as being proportional. The eigenvalue equation

$$(H_0 + H_1)|\ \rangle = H'|\ \rangle \quad \text{or} \quad (H_0 - H')|\ \rangle = -H_1|\ \rangle, \qquad (10.70)$$

then breaks up into the series of equations

$$(H_0 - H_0')|H_0'\rangle = 0, \qquad (10.71.0)$$

$$(H_0 - H_0')|\ \rangle_1 - E_1|H_0'\rangle = -H_1|H_0'\rangle \qquad (10.71.1)$$

$$(H_0 - H_0')|\ \rangle_2 - E_1|\ \rangle_1 - E_2|H_0'\rangle = -H_1|\ \rangle_2, \qquad (10.71.2)$$

$$(H_0 - H_0')|\rangle_3 - E_1|\rangle_2 - E_2|\rangle_1 - E_3|H_0'\rangle = -H_{1S}|\rangle_2 \quad (10.71.3)$$

. . . . . .

for the solutions of which recurrence formulae can be established.

Multiplying equation (10.71.1) on the left by $\langle H_0'|$, one obtains for the first-order correction to the energy eigenvalues (cf. (6.53))

$$E_1 = \langle H_0'|H_{1S}|H_0'\rangle \quad (10.72)$$

Equation (10.71.1) may also be solved formally to express in terms of $|H_0'\rangle$

$$|\rangle_1 = -\frac{1}{H_0 - H_0'}(H_{1S} - E_1)|H_0'\rangle \quad (10.73)$$

The operator $\frac{1}{H_0 - H_0'}$ evidently has a pole at $H_0 = H_0'$. In (10.73), however, this pole is cancelled by the zero of the operator $H_{1S} - E$ via (10.72). In fact we assume that the cancellation is of such a nature that

$$\langle H_0'|\rangle_1 = -\frac{1}{H_0' - H_0}\langle H_0'|H_{1S} - E_1|H_0'\rangle = 0 \quad (10.74)$$

Multiplying equation (10.71.2) on the left by $\langle H_0'|$ and making use of (10.74), one obtains for the second order correction to the energy eigenvalues

$$E_2 = \langle H_0'|H_{1S}|\rangle_1 \quad (10.75)$$

Inserting (10.73) into (10.75), we may write explicitly

$$E_2 = -\langle H_0'|H_{1S}\frac{1}{H_0 - H_0'}(H_{1S} - E_1)|H_0'\rangle \quad (10.76)$$

In equation (10.76), the factor $\frac{1}{H_0 - H_0'}(H_1 - E_1)|H_0'\rangle$ may also be written in the form $\frac{1_0'}{H_0 - H_0'} H_1 |H_0'\rangle$ where the operator $1_0'$ is defined by

$$1_0' \equiv 1 - |H_0'\rangle\langle H_0'| \tag{10.77}$$

Hence, (10.76) becomes (cf. (6.58))

$$\begin{aligned} E_2 &= -\langle H_0' | H_1 \frac{1_0'}{H_0 - H_0'} H_1 | H_0' \rangle \\ &= -\sum_{H_0''}{}' \langle H_0' | H_1 | H_0'' \rangle \frac{1}{H_0'' - H_0'} \langle H_0'' | H_1 | H_0' \rangle \\ &= -\sum_{H_0''}{}' \frac{|\langle H_0'' | H_1 | H_0' \rangle|^2}{H_0'' - H_0'} \end{aligned} \tag{10.78}$$

where the prime indicates a summation over all states except $|H_0'\rangle$ Since the operators $1_0'$ and H_0 commute, the notation $\frac{1_0'}{H_0 - H_0'}$ is unambiguous.

Solving eq. (10.71.2) formally, we obtain

$$|\rangle_2 = \frac{1}{H_0 - H_0'} \left\{ -(H_1 - E_1)|\rangle_1 + E_2 |H_0'\rangle \right\} \tag{10.79}$$

Again we assume that the pole of the operator $\frac{1}{H_0 - H_0'}$ is cancelled, in virtue of (10.74) and (10.75) in such a manner that

$$\langle H_0' | \rangle_2 = 0 \tag{10.80}$$

Continuing in this fashion we obtain the series of equations

$$\begin{aligned} E_1 &= \langle H_0' | H_1 | H_0' \rangle \\ E_2 &= \langle H_0' | H_1 | \rangle_1 \\ E_3 &= \langle H_0' | H_1 | \rangle_2 \\ E_4 &\quad \langle H_0' | H_1 | \rangle_3 \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{10.81}$$

and

$$|\rangle_1 = -\frac{1}{H_0 - H_0'}(H_{1S} - E_1)|H_0'\rangle$$

$$|\rangle_2 = -\frac{1}{H_0 - H_0'}(H_{1S} - E_1)|\rangle_1 + \frac{1}{H_0 - H_0'}E_2|H_0'\rangle \quad (10.82)$$

$$|\rangle_3 = -\frac{1}{H_0 - H_0'}(H_{1S} - E_1)|\rangle_2 + \frac{1}{H_0 - H_0'}E_2|\rangle_1 + \frac{1}{H_0 - H_0'}E_3|H_0'\rangle$$

. . . . . .

Problem LI : Show that

$$E_3 = \sum_{H_0'' H_0'''}{}' \langle H_0'|H_{1S}|H_0''\rangle \frac{1}{H_0'' - H_0'} \langle H_0''|H_{1S} - E_1|H_0'''\rangle \frac{1}{H_0''' - H_0'} \langle H_0'''|H_{1S}|H_0$$

and

$$E_4 = -\sum_{H_0'' H_0''' H_0''''}{}' \langle H_0'|H_{1S}|H_0''\rangle \frac{1}{H_0'' - H_0'} \langle H_0''|H_{1S} - E_1|H_0'''\rangle \frac{1}{H_0''' - H_0'} \langle H_0'''|H_{1S} - E_1|H_0$$

$$\cdot \frac{1}{H_0'''' - H_0'} \langle H_0''''|H_{1S}|H_0'\rangle - E_2 \sum_{H_0''}{}' \langle H_0'|H_{1S}|H_0''\rangle \frac{1}{(H_0'' - H_0')^2} \langle H_0''|H_{1S}|H_0$$

General recurrence formulae for the E's can be found in Born Heisenberg, Jordan, Zeitschrift für Physik, 35, p. 565 (1925)

If we add equations (10.81) together, we get

$$\Delta E = \langle H_0'|H_{1S}|\rangle \quad (10.83)$$

where $\Delta E = E_1 + E_2 + E_3 + \cdots = H' - H_0'$

(10.84)

Furthermore, if we add equations (10.82) together, we get

$$|\rangle - |H_0'\rangle = -\frac{1}{H_0 - H_0'}(H_{1S} - \Delta E)|\rangle \tag{10.85}$$

In addition to the series of equations (10.81) and (10.82), we must also append the equations

$$\begin{aligned} \langle H_0' | \rangle_1 &= 0 \\ \langle H_0' | \rangle_2 &= 0 \\ \langle H_0' | \rangle_3 &= 0 \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{10.86}$$

which, when added together, give

$$\langle H_0' | (|\rangle - |H_0'\rangle) = 0 \tag{10.87}$$

or

$$\langle H_0' | \rangle = \langle H_0' | H_0' \rangle = 1 \tag{10.88}$$

Using (10.83) and (10.88), we may write

$$\begin{aligned} \mathcal{T}_0'(H_{1S} - \Delta E)|\rangle &= (H_{1S} - \Delta E)|\rangle - |H_0'\rangle\langle H_0'|H_{1S} - \Delta E|\rangle \\ &= (H_{1S} - \Delta E)|\rangle - |H_0'\rangle(\Delta E - \Delta E) \\ &= (H_{1S} - \Delta E)|\rangle \end{aligned} \tag{10.89}$$

Hence equation (10.85) may be rewritten in the following convenient form,

$$|\rangle = |H_0'\rangle - \frac{1_0'}{H_0 - H_0'} (H_{1s} - \Delta E)\,|\rangle \qquad (10.90)$$

which has the formal solution

$$|\rangle = \left[1 + \frac{1_0'}{H_0 - H_0'} (H_{1s} - \Delta E)\right]^{-1} |H_0'\rangle \qquad (10.91)$$

Combining eqs. (10.83) and (10.91) we have the following equation for the level shift ΔE :

$$\Delta E = \langle H_0' | H_{1s} \left[1 + \frac{1_0'}{H_0 - H_0'} (H_{1s} - \Delta E)\right]^{-1} |H_0'\rangle \qquad (10.92)$$

Or, making use also of (10.88), we may write

$$0 = \langle H_0' | (H_{1s} - \Delta E)\left[1 + \frac{1_0'}{H_0 - H_0'} (H_{1s} - \Delta E)\right]^{-1} |H_0'\rangle \qquad (10.93)$$

If the operator $\left[1 + \frac{1_0'}{H_0 - H_0'} (H_{1s} - \Delta E)\right]^{-1}$ is expanded by the usual binomial theorem, then the right hand side of (10.93) becomes a power series in ΔE . The problem of determining the level shift ΔE becomes one of finding the zero of this power series. If the power series has more than one zero then it is necessary to choose that one for which $\Delta E \to 0$ as $H_1 \to 0$. In practice one terminates the series after a few terms, leaving a polynomial in ΔE. In simple cases this polynomial could be plotted and its zero determined graphically. The legitimacy of terminating the series, like the whole question of the convergence of the expansions for the individual coefficients, depends on the smallness of H_1 .

The above procedure can readily be generalized to include the case of degeneracy. Instead of starting from the old fashioned approach, let us use modern formal methods from the outset. Since the system is now degenerate, we shall need some extra labels γ , in addition to H_0^1, in order to specify the quantum states. Let us try to write the eigenvector of H in the form

$$|\rangle = |\rangle_{H_0'} + |\rangle_{\Delta} \tag{10.94}$$

where

$$\langle H_0', \gamma' | \rangle_{\Delta} = 0 \quad \text{for all } \gamma' \tag{10.95}$$

and where the subscript H_0' on the vector $|\rangle_{H_0'}$ indicates that $|\rangle_{H_0'}$ is an eigenvector of H_0 corresponding to the eigenvalue H_0', but that we can say nothing at present about its specification in terms of the labels γ' . Then, writing

$$H' = H_0' + \Delta E \tag{10.96}$$

we must have

$$\begin{aligned} H_0' |\rangle_{H_0'} + H_0' |\rangle_{\Delta} + \Delta E |\rangle &= H' |\rangle = H |\rangle \\ &= H_0 |\rangle_{H_0'} + H_0 |\rangle_{\Delta} + H_{1S} |\rangle \\ &= H_0' |\rangle_{H_0'} + H_0 |\rangle_{\Delta} + H_{1S} |\rangle \end{aligned} \tag{10.97}$$

or

$$(H_0 - H_0') |\rangle_{\Delta} = -(H_{1S} - \Delta E) |\rangle \tag{10.98}$$

or

$$|\rangle = |\rangle_{H_0'} - \frac{1}{H_0 - H_0'} (H_{1S} - \Delta E) |\rangle \tag{10.99}$$

Multiplying eq. (10/98) on the left by $\langle H_0', \gamma' |$ and making use of (10.95), we obtain

$$\Delta E \langle H_0', \gamma' | \rangle_{H_0'} = \langle H_0', \gamma' | H_{1S} | \rangle \tag{10.100}$$

If we generalize equation (10.77) by the definition

$$1'_0 \equiv 1 - \sum_{\gamma'}{}' |H_0', \gamma'\rangle\langle H_0'\gamma'|, \tag{10.101}$$

we may also write

$$1'_0 (H_{1S} - \Delta E)|\rangle = (H_{1S} - \Delta E)|\rangle \tag{10.102}$$

Hence, equation (10.99) may be rewritten in the form

$$|\rangle = |\rangle_{H_0'} - \frac{1_0'}{H_0 - H_0'} (H_{1S} - \Delta E)|\rangle \tag{10.103}$$

of which the formal solution is

$$|\rangle = \left[1 - \frac{1_0'}{H_0 - H_0'} (H_{1S} - \Delta E)\right]^{-1} |\rangle_{H_0'} \tag{10.104}$$

Substituting (10.104) into (10.100), we obtain

$$\Delta E \langle H_0', \gamma'|\rangle_{H_0'} = \langle H_0', \gamma'| H_{1S} \left[1 - \frac{1_0'}{H_0 - H_0'} (H_{1S} - \Delta E)\right]^{-1} |\rangle_{H_0'} \tag{10.105}$$

In virtue of the equation

$$\langle H_0', \gamma' | H_{1S} - \Delta E |\rangle = 0 \tag{10.106}$$

(10.105) may also be rewritten in the form

$$\begin{aligned} 0 &= \langle H_0', \gamma' | (H_{1S} - \Delta E) \left[1 - \frac{1_0'}{H_0 - H_0'} (H_{1S} - \Delta E)\right]^{-1} |\rangle_{H_0'} \\ &= \sum_{\gamma''}{}' \langle H_0, \gamma' | (H_{1S} - \Delta E) \left[1 + \frac{1_0'}{H_0 - H_0'} (H_{1S} - \Delta E)\right]^{-1} \\ &\qquad |H_0', \gamma''\rangle\langle H_0', \gamma''|\rangle_{H_0'} \end{aligned} \tag{10.107}$$

Equation (10.107) says that, for each value of H_0', the matrix

$$\left(\langle H_0', \gamma' | (H_1 - \Delta E) \left[1 + \frac{1}{H_0 - H_0'} (H_1 - \Delta E) \right]^{-1} | H_0', \gamma'' \rangle \right) \qquad (10.108)$$

must be singular. To obtain the possible values of ΔE for each H'_0 one therefore takes the determinant of the matrix (10.108) and sets it equal to zero. The result is an equation in ΔE, whose roots are the desired values. The equation will in general have as many roots as there are values of γ' for a given H'_0. If all the roots are distinct the perturbation is said to have completely removed the degeneracy from the original energy level. If some of the roots coincide, then one says that the perturbation has only partially removed the degeneracy.

In practice one makes an approximate evaluation of the determinant of (10.108) by expanding the operator (10.108) $\left[1 + \frac{1}{H_0 - H_0'} (H_1 - \Delta E) \right]^{-1}$ by the binomial theorem and terminating the series after a few terms. The result is a polynomial in ΔE. The level shifts ΔE are then given by those zeros of this polynomial which vanish in the limit $H_1 \longrightarrow 0$. Once these shifts are determined one can substitute the values into (10.108) and choose the 1-column matrices corresponding to the eigenvalue 0. The normalization condition

$$\sum_{\gamma''}' | \langle H_0', \gamma'' | \rangle_{H_0'} |^2 = 1 \qquad (10.109)$$

should also be imposed.

By a suitable choice of the labels γ' the matrices (10.108), for the permissible values of ΔE, may be made diagonal. When the γ''s have been chosen in this fashion we shall denote the complete set of labels H'_0, simply by α'_0. Equations (10.103, 104, 106, 107) may then be written in the forms

$$| \rangle = | \alpha_0' \rangle - \frac{1}{H_0 - H_0'} (H_1 - \Delta E) | \rangle \qquad (10.110)$$

$$|> = \left[1 + \frac{1_0'}{H_0 - H_0'}(H_{1S} - \Delta E)\right]^{-1} |\alpha_0'> \qquad (10.111)$$

$$\langle \alpha_0' | H_{1S} - \Delta E | > = 0 \qquad (10.112)$$

In the continuous case, when there is no level shift, we have H' = H_0' and therefore

$$\langle \alpha_0' | H | >_{\alpha_0'} = H' \langle \alpha_0' | >_{\alpha_0'} = H_0' \langle \alpha_0' | >_{\alpha_0'}$$

Also $$\langle \alpha_0' | H_0 | >_{\alpha_0'} = H_0' \langle \alpha_0' | >_{\alpha_0'}$$

Substracting one equation from tjo other we get

$$\langle \alpha_0' | H_{1S} | >_{\alpha_0'} = 0$$

which is the analog of (10.112)

$$0 = \langle \alpha_0' | (H_{1S} - \Delta E) \left[1 + \frac{1_0'}{H_0 - H_0'}(H_{1S} - \Delta E)\right]^{-1} |\alpha_0'> \qquad (10.113)$$

We also have the equation (cf. (10.88))

$$\langle \alpha_0' | > = 1 \qquad (10.114)$$

provided $|\alpha_0'>$ is normalized.

Let us now return to the general theory. It will immediately be observed that equation (10.110) looks very much like the integral equation (10.63), but with H_{1S} replaced by $H_{1S} - \Delta E$. Suppose the perturbation function had originally been taken to be $H_1 - \Delta E$ instead of H_1 . Then if we had started from $|\alpha_0'>$ as the original unperturbed state, the energy level would had suffered no shift in passing to the perturbed state $|>$. This causes us to hope that by a suitable readjustement of the definition of the perturbation function in each case we may, after all, be able to obtain the perturbed state through appli-

cation of the operator $U_I'(t_0, -\infty)$.

The situation is not actually as simple as this, but the idea is suggestive. Denote by $U_I'(t_0, -\infty)$ the operator obtained through use of $H_{1,S} - \Delta E$ instead of $H_{1,S}$. Then applying formula (10.25) and inserting a damping factor after the manner of (10.46), we obtain

$$U_I'(t_0, -\infty) = T\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0}(H_{1,I}(t) - \Delta E)\, e^{\frac{\varepsilon}{\hbar}(t-t_0)}\, dt\right]$$

$$= T\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} H_{1,I}(t)\, e^{\frac{\varepsilon}{\hbar}(t-t_0)}\, dt\right]\exp\left(-\frac{\Delta E}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{\varepsilon}{\hbar}(t-t_0)}\, dt\right)$$

$$= \frac{U_I(t_0, -\infty)}{\frac{\Delta E}{i\varepsilon}} \qquad (10.115)$$

Neither the numerator nor the denominator in the last expression is well-defined in the limit $\varepsilon \to 0$. The operator $U_I'(t_0, -\infty)$ is here, however, represented as the ratio of the two quantities. This suggests that we may represent the true perturbed state-vector $|\ \rangle$ as a convergent ratio of two convergent quantities. Accordingly, we shall set

$$|\ \rangle = \frac{U_I(t_0, -\infty)\,|\alpha_0'\rangle}{\langle\alpha_0'|U_I(t_0, -\infty)|\alpha_0'\rangle} = \frac{|\ \rangle_{\alpha_0'}}{\langle\alpha_0'|\ \rangle_{\alpha_0'}} \qquad (10.116)$$

where $|\ \rangle_{\alpha_0'}$ is given by (10.62). We see first of all that the vector $|\ \rangle$ thus defined satisfies equation (10.114). Let us now examine some of its analytical properties. Using formula (10.25) and explicitly exhibiting the damping factors, we have

$$|\ \rangle_{\alpha_0'} = \sum_{n=0}^{\infty}{}' \frac{1}{n!\,(i\hbar)^n}\int_{-\infty}^{t_0} dt_1 \cdots \int_{-\infty}^{t_0} dt_n\, e^{\frac{\varepsilon}{\hbar}(t_1 + \cdots + t_n - n t_0)} \times$$

$$T\left[H_{1,I}(t_1)\cdots H_{1,I}(t_n)\right]|\alpha_0'\rangle \qquad (10.117)$$

Then, remembering that $\dot{F}_I = \frac{1}{i\hbar}[F_I, H_0]$ for operators in the interaction representation which have no explicit dependence on the

$$* \quad (\ )_n = (t_1 + \cdots + t_n - n t_0)$$

time, we may write

$$(H_0 - H_0')|\rangle_{\alpha_0'} = H_0 U_I(t_0, -\infty)|\alpha_0'\rangle - U_I(t_0, -\infty) H_0'|\alpha_0'\rangle$$

$$= -[U_I(t_0, -\infty), H_0]|\alpha_0'\rangle$$

$$= -\sum_{n=1}^{\infty}{}' \frac{1}{n!(ik)^n} \int_{-\infty}^{t_0} dt_1 \cdots \int_{-\infty}^{t_0} dt_n \, e^{\frac{\varepsilon}{k}(t_1 + \cdots + t_n - n t_0)} \cdot \sum_{i=1}^{n}{}' \frac{\partial}{\partial t_i} P[H_{1I}(t_1) \cdots H_{1I}(t_n)]|\alpha_0'\rangle$$

$$= \sum_{n=1}^{\infty}{}' \frac{1}{(n-1)!(ik)^n} \int_{-\infty}^{t_0} dt_1 \cdots \int_{-\infty}^{t_0} dt_n \, e^{\frac{\varepsilon}{k}(\ )_n} \frac{\partial}{\partial t_n} P[H_{1I}(t_1) \cdots H_{1I}(t_n)]|\alpha_0'\rangle$$

$$= -\sum_{n=1}^{\infty}{}' \frac{1}{(n-1)!(ik)^n} \int_{-\infty}^{t_0} dt_1 \cdots \int_{-\infty}^{t_0} dt_{n-1} \, e^{\frac{\varepsilon}{k}(\ )_{n-1}} * H_{1I}(t_0) P[H_{1I}(t_1) \cdots H_{1I}(t_{n-1})]|\alpha_0'\rangle$$

$$+ \frac{\varepsilon}{k} \sum_{n=1}^{\infty}{}' \frac{1}{(n-1)!(ik)^n} \int_{-\infty}^{t_0} dt_1 \cdots \int_{-\infty}^{t_0} dt_n \, e^{\frac{\varepsilon}{k}(\ )_n} P[H_{1I}(t_1) \cdots H_{1I}(t_n)] H_0'|\rangle$$

$$= -H_{1I}(t_0) U_I(t_0, -\infty)|\alpha_0'\rangle + i\varepsilon\lambda \frac{\partial}{\partial\lambda} U_I(t_0, -\infty)|\alpha_0'\rangle$$

$$= -H_{1I}|\rangle_{\alpha_0'} + i\varepsilon\lambda \frac{\partial}{\partial\lambda}|\rangle_{\alpha_0'} \qquad (10.118)$$

where we have used the fact that $H_{1I}(t_0) = H_{1S}$ and where we have imagined the perturbation function H_1 to be proportional to a parameter λ with respect to which we have differentiated in the last two lines. In a field theory λ would be known as a "coupling constant".

We may now write

$$i\varepsilon\lambda \frac{\partial}{\partial\lambda}|\rangle = i\varepsilon\lambda \left[\frac{\frac{\partial}{\partial\lambda}|\rangle_{\alpha_0'}}{\langle\alpha_0'|\rangle_{\alpha_0'}} - |\rangle_{\alpha_0'} \frac{\frac{\partial}{\partial\lambda}\langle\alpha_0'|\rangle_{\alpha_0'}}{\langle\alpha_0'|\rangle_{\alpha_0'}^2} \right]$$

$$= \frac{(H_0 + H_{1S} - H_0')|\rangle_{\alpha_0'}}{\langle\alpha_0'|\rangle_{\alpha_0'}} - i\varepsilon\lambda \frac{\partial}{\partial\lambda} \log \langle\alpha_0'|\rangle_{\alpha_0'} \cdot \frac{|\rangle_{\alpha_0'}}{\langle\alpha_0'|\rangle_{\alpha_0'}}$$

$$= \left(H - H_0' - i\varepsilon \lambda \frac{\partial}{\partial \lambda} \log \langle \alpha_0' | \rangle_{\alpha_0'} \right) | \rangle \qquad (10.119)$$

Since we can be reasonablu confident that the vector $| \rangle$ is convergent (although we have not proved it yet), the left hand side of (10.119) vanishes in the limit $\varepsilon \to 0$, and we have

$$H | \rangle = \left(H_0' - i\varepsilon \lambda \frac{\partial}{\partial \lambda} \log \langle \alpha_0' | \rangle_{\alpha_0'} \right) | \rangle \qquad (10.120)$$

The level shift may evidently be identified as

$$\Delta E = \lim_{\varepsilon \to +0} i\varepsilon \lambda \frac{\partial}{\partial \lambda} \log \langle \alpha_0' | \rangle_{\alpha_0'} \qquad (10.121)$$

Multiplying equation (10.120) on the left by $\langle \alpha_0' |$, we obtain

$$\langle \alpha_0' | H | \rangle = \langle \alpha_0' | H_0 - H_{0s} | \rangle = \langle \alpha_0' | H_0' + H_{1s} | \rangle$$

$$= \langle \alpha_0' | H_0' + \Delta E | \rangle \qquad (10.122)$$

which leads to (10.112).

The final confirmation of the fact that $| \rangle$ is a well defined eigenvector of H may be obtained by actually calculating $| \rangle_{\alpha_0'}$, and $\langle \alpha_0' | \rangle$, substituting in (10.116) and passing to the limit $\varepsilon \to 0$. We shall verify this up to the first order of approximation :

Using (10.48), we have, to first order

$$| \rangle_{\alpha_0'} = | \alpha_0' \rangle - \frac{1}{H_0 - H_0' - i\varepsilon} H_{1s} | \alpha_0' \rangle + \cdots \qquad (10.123)$$

and hence

$$\langle \alpha_0' | \rangle_{\alpha_0'} = 1 - \frac{1}{i\varepsilon} \langle \alpha_0' | H_{1s} | \alpha_0' \rangle + \cdots \qquad (10.124)$$

Thus

$$|\rangle = \lim_{\varepsilon \to 0} \frac{|\alpha_0'\rangle - \frac{1}{H_0' - H_0' - i\varepsilon} H_1 |\alpha_0'\rangle + \cdots}{1 - \frac{1}{i\varepsilon} \langle \alpha_0' | H_1 | \alpha_0' \rangle + \cdots}$$

$$= \lim_{\varepsilon \to 0} \left\{ |\alpha_0'\rangle - \left(\frac{1}{H_0 - H_0' + i\varepsilon} H_1 - \frac{1}{i\varepsilon} \langle \alpha_0' | H_1 | \alpha_0' \rangle \right) |\alpha_0'\rangle + \cdots \right\}$$

$$= \lim_{\varepsilon \to 0} \left\{ |\alpha_0\rangle - \frac{1}{H_0 - H_0' + i\varepsilon} \left(H_1 - \langle \alpha_0' | H_1 | \alpha_0' \rangle \right) |\alpha_0'\rangle + \cdots \right\} \qquad (10.125)$$

Remembering that $\langle \alpha_0' | H_1 | \alpha_0' \rangle$ is the first approximation to ΔE we see that (10.125) is an explicit expansion of equation (10.111). Expressions (10.111) and (10.116) therefore define the same vector. One must always remember, of course, that the labels α_0' must be chosen appropriately so that the removal of any degeneracy by the perturbation function H_1 is a diagonal process. If the lables α_0' are not chosen appropriately, the vector $|\rangle$ will not converge even when defined by (10.116).

The vector $|\rangle$ will, in general, not be normalized when constructed according to the recipe (10.111) or (10.116). This can be seen by writing $|\rangle$ in the form

$$|\rangle = |\alpha_0'\rangle + |\rangle_\Delta \qquad (10.126)$$

and remembering the condition

$$\langle \alpha_0' | \rangle_\Delta = 0 \qquad (10.127)$$

We then have

$$\langle 1 \rangle = \langle \alpha_0 | \alpha_0 \rangle + \langle 1 \rangle_\Delta$$

$$= 1 + \langle 1 \rangle_\Delta > 1 \qquad (10.128)$$

It is therefore necessary to normalize $|\ \rangle$ after having constructed it by (10.111) or (10.116).

This lack of normality of $|\ \rangle$ may perhaps be surprising if one looks only at the recipe (10.111) for constructing it. For, (10.111) results directly from the integral equation for the operator $U_I'(t_0, -\infty)$ of (10.115), in the limit $\varepsilon \to 0$. As long as ε remains finite $U_I'(t_0, -\infty)$ is unitary, and whatever else it does when it acts upon the state-vector $|\alpha_0\rangle$ it does not alter the normalization: Something peculiar must happen, however in the limit $\varepsilon \to 0$, and this must be even more true for the operator $U_I(t_0, -\infty)$. That is, associated with the divergence of the operator $U_I(t_0, -\infty)$ there will in general be a certain lack of unitarity, in the limit $\varepsilon \to 0$. The possibility of such behavior may be illustrated by the following considerations. The operator $U_I(t, -\infty)$ satisfies the differential equation (cf. (10.20))

$$i\hbar \frac{d}{dt} U_I(t, -\infty) = H_{1I}(t)\, U_I(t, -\infty). \qquad (10.129)$$

By a procedure analogous to that of eqs. (8.28) we may write

$$\frac{d}{dt}\left[U_I^*(t, -\infty) U_I(t, -\infty)\right] = \frac{1}{i\hbar} U_I^*(t, -\infty)\left[-H_{1I}(t) + H_{1I}(t)\right] U_I(t, -\infty) = 0 \qquad (10.130)$$

From (10.58) we have

$$\lim_{t \to -\infty} U_I^*(t, -\infty)\, U_I(t, -\infty) = 1, \qquad (10.131)$$

Hence the inference is that

$$U_I^*(t, -\infty)\, U_I(t, -\infty) = 1 \quad \text{for all } t \qquad (10.132)$$

On the other hand, since the lower end of the time interval is already at $-\infty$, we cannot make use of (10.21) and the analog of eqs. (8.30, 31), and hence we cannot infer that $U_I(t,-\infty)\,U_I^*(t,-\infty) = 1$. The most that we can show, using (10.132), is that the product $U_I(t,-\infty)\,U_I^*(t,-\infty)$ is idempotent :

$$\left[U_I(t,-\infty)\,U_I^*(t,-\infty)\right]^2 = U_I(t,-\infty)\,U_I^*(t,-\infty) \quad (10.133)$$

The eigenvalues of the operator $U_I(t,-\infty)\,U_I^*(t,-\infty)$ may therefore be divergent or else equal to 0 as well as to 1. Consequently we must be prepared to find that $U_I(t,-\infty)$ often has only a sort of "quasi unitary" property rather than being truly unitary. Stated in another way, we must be prepared to find that the unitarity of $U_I(t,-\infty)$, in cases in which it is non-convergent, can be maintained only by interpreting its non-convergent matrix elements in very special ways.

Let us now return to a consideration of those cases in which the perturbation does not cause a shift in the energy levels and in which the operator $U_I(t,-\infty)$ is perfectly well behaved. This is the typical situation in scattering problems where there are no bound states. (The case of bound states will be discussed later). In such problems a particle is imagined to be started off initially in an essentially "free" condition at a large distance from a scattering center but moving toward it. At a much later time the particle has been scattered and is moving away from the scattering center, again in an essentially "free" condition. Now, the energy eigenvalue for the state in which the particle happens to be during this process can be determined from the asymptotic part of the stationary-state wave function of the particle at a large distance from the scattering center just as well as from the wave function at points closer to it. But at large distances the wave function is essentially that of a free particle. Hence the energy of the perturbed (i.e. scattered) particle is the same as that of the free particle, and there is no shift in the energy levels.

In scattering problems the perturbation also will in general not remove any degeneracy which the unperturbed system may posses. Hence we may ignore the question of degeneracy.

A typical feature of scattering systems, or, more generally, of systems for which the operator $U_I(t,-\infty)$ is well behaved, is that the energy eigenvalues form a continuum. It must not be supposed, however, that this condition is sufficient to insure respectable behavior of $U_I(t,-\infty)$. One may, for example, imagine perturbed systems which possess a continuum of energy eigenvalues but for which the perturbation is extensive enough so that the entire continuum suffers a level shift. This situation occurs, for example, in the quantum theory of fields, the shift in the level-continuum being produced by the "self-energy" effect. It is possible also for the perturbation to effect a total or partial removal of degeneracy even from a continuum of levels, although this will generally not happen in the case of field theories, since, there, the degeneracy is connected with the invariant transformation properties of the field system, and there are not altered by coupling with other field systems.

In order to insure that the operator $U_I(\ )$ will be well behaved we must therefore assume not only that the energy eigenvalues form a continuum but also that the perturbation produces no level shift and removes no degeneracy from the unperturbed system. Under these circumstances it would appear that an unperturbed energy eigenstate is transformed in a straightforward manner into a perturbed energy eigenstate through an adiabatic application of the perturbation H_I. We can, in fact, show that if $|\psi_0'\rangle$ is the Schrödinger state-vector of the system at "switch-on" time (i.e. beginning of the adiabatic build-up process) then the Schrödinger state-vector is, at time t_0, given by $|\ \rangle_{\psi_0'}$ apart from a phase factor. For, using (10.9) and (10.26), we have

$$\lim_{t\to-\infty} |t\rangle = \lim_{t\to-\infty} e^{\frac{i}{\hbar}H_0(t_0-t)}\, U_I(t,-\infty)\, |t\rangle \qquad (10.134)$$
$$= e^{i\varphi}\,|\psi_0'\rangle$$

where φ is t e phase difference between the Schrödinger and interaction state-vectors at "switch-on" time. Hence

$$|t_0\rangle = |\rangle = |t_0\rangle_I = U_I(t_0,t)|t\rangle_I$$
$$= \lim_{t\to-\infty} U_I(t_0,t)|\rangle_I = e^{i\varphi} U_I(t_0,-\infty)|\alpha_i'\rangle \quad (10.135)$$
$$= e^{i\varphi}|\rangle_{\alpha_i'}$$

Actually, even in the case in which the energy eigenvalues form a discrete set the adiabatic application of the perturbation will transform an unperturbed state into the corresponding perturbed state. The only trouble in this case is that the transformation operator $U_I(t_0,-\infty)$ becomes undefined in the limit $\epsilon \to 0$. We get around this difficulty by defining the perturbed state $|\rangle$ according to the formula (10.116). Thus (10.116) is an explicit statement of the quantum analog of the classical theorem of adiabatic invariance which was briefly discussed in chaptor 2. There is a one-to-one correspondence between the unperturbed states and the perturbed states. They may therefore be given the same numerical labels -a process which is justified in the old quantum theory by the constency in the numerical values of tho instantaneous action variables.

When the operator $U_I(t_0,-\infty)$ is well behaved it becomes once more of interest to study the integral equation (10.61) in some detail. Let us write

$$U_I(t_0,-\infty) = 1 + T_- \quad (10.136)$$

where

$$\langle\alpha_a''|T_-|\alpha_b'\rangle = \frac{1}{\hbar}(W_a'' - W_b')\langle\alpha_a''|\chi_-|\alpha_b'\rangle \quad (10.137)$$

with

$$\chi_- = -2\pi i \, H_{1S} \, U_I(t_0 = -\infty) \qquad (10.38)$$

Now, when the operator $U_I(t_0 = -\infty)$ is applied to the stationary state vector $|\alpha_0'\rangle$ of the unperturbed system, we get the stationary state vector $|\ \rangle_{\alpha_0'}$ of the perturbed system - and we get it complete in every detail. For example; if we are concerned with a scattering problem we can obtain from $|\ \rangle_{\alpha_0'}$ detailed information about the structure of the wave-function everywhere, even at points close in to the scattering center. Suppose, however, we are not interested in the details of the wave-function close to the scattering center but only in its asymptotic behavior at large distances. We know that the asymptotic wave function behaves like a free wave function corresponding to the same energy as the unperturbed wave function. Hence for the study of the asymptotic form of the wave function we are only interested in matrix elements $\langle \alpha_0''|\ \rangle_{\alpha_0'}$ for which $H_0'' = H_0'$. More precisely, we are only interested in the matrix elements.

$$= \frac{i}{2}\langle \alpha_0''|\underline{\chi}_-|\alpha_0'\rangle \Rightarrow$$

$$\delta_{H_0''H_0'}\langle \alpha_0''|T_-|\alpha_0'\rangle = \frac{i}{2}\delta(H_0''-H_0')\langle \alpha_0''|\chi_-|\alpha_0'\rangle \qquad (10.139)$$

where the notation involving a bar underlining an operator is defined by

$$\langle \alpha_0''|\underline{F}|\alpha_0'\rangle \equiv \delta(H_0''-H_0')\langle \alpha_0''|F|\alpha_0'\rangle \qquad (10.140)$$

Underlined operators only have non-vanishing matrix elements connecting states of equal energy. It is evident that operators of this type are more suitable for the discussion of asymptotic behavior than is the operator $U_I(t_0 = -\infty)$ itself. In particular, we see that for a quantitative discussion of asymptotic behavior, an operator of the form $1 + \alpha\,\underline{\chi}_-$ where α is a certain constant, will play a very important role. We cannot say from the foregoing simple discussion just what the value of α should be. For, although we can be certain that

the projection $|\alpha'\rangle\langle\alpha'|\,\psi_{\alpha''}\rangle$ of $|\psi_{\alpha''}\rangle$ on $|\alpha'\rangle$ "scops up" some of the state vector $|\psi_{\alpha''}\rangle$ we cannot say for sure just how much of $|\psi_{\alpha''}\rangle$ it "scoops up", nor how much weight this portion shall be given in comparison with projections on other unperturbed state-vectors with the same energy. (*)

We can determine the value of a by wave packet argument. Suppose the time dependent Schrödinger wave function of the system has the form of a plane wave packet coming in from infinity. This wave packet can be built out of stationary state wave functions of the perturbed system, together with appropriate exponential energy-time factors $e^{-\frac{i}{\hbar}Et}$ However, while it is still at a large distance from the scattering center it can equally well be built out of the corresponding stationary state wave functions of the unperturbed system. Moreover, while it is at a large distance it doesn't very much matter whether the scattering perturbation H_1 is present or not. Hence H_1 can be switched on adiabatically while the wave packet approaches. When the packet gets near to the scattering center then it becomes necessary to use the perturbed wave functions. However, after the packet has been scattered (it generally gets broken up in the process) ans is moving away from the scattering center then the use of the unperturbed wave functions can be resumed. The unperturbed wave functions used after the scattering process will not be those used prior to the scattering process, but will be connected (apart from phase factors arising from the difference between the Schrödinger and interaction representations) with the previous wave functions by the operator

$$S = U_{\infty}\,(\infty, -\infty) \qquad (10.141)$$

On the other hand, the perturbed wave functions out of which the packet can be constructed must remain fixed for all time. Therefore, if we pass to the limit of a very broad packet with a very narrow (effectively δ-function) incoming momentum distribution, the corresponding perturbed wave function reduces effectively to a single stationary state wave func-

(*) For a more complete discussion of the argument developed hereafter (up to (10.142) page 38), see Appendix page 77.

tion (with appropriate exponential energy-time factor), and its asymptotic incoming and outgoing parts must be connected by the operator above*.

The operator S is known as the scattering operator, and in its matrix form, is often referred to as the S-matrix. It is the operator which we are seeking above as giving the correct quantitati description of asymptotic behavior. From (10.45), (10.56) and (10.60) we find that it is given by

$$S = 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} H_{1I}(t')\, U_I(t',-\infty)\, dt'$$

$$= 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} e^{-\frac{1}{i\hbar}H_0(t'-t_0)} H_{1S}\, U_I(t_0,-\infty)\, e^{\frac{1}{i\hbar}H_0(t'-t_0)}\, dt'$$

$$= 1 + \frac{1}{i\hbar}\sum_{\alpha_0'}\int_{-\infty}^{\infty} e^{-\frac{1}{i\hbar}(H_0-H_0')t}\, dt\, H_{1S}\, U_I(t_0,-\infty)\,|\alpha_0'\rangle\langle\alpha_0'|$$

$$= 1 - 2\pi i \sum_{\alpha_0'} \delta(H_0-H_0')\, H_{1S}\, U_I(t_0,-\infty)\,|\alpha_0'\rangle\langle\alpha_0'|$$

$$= 1 + \underline{K}_- . \qquad (10.142)$$

The constant a mentionned above is thus seen to be equal to unity.

Although it is not desirable to carry out the mathematical discussion of the S-matrix in terms of a series expansion; for computational purposes such expansions are often useful. Combining eqs. (10.57) and (10.142), we may write :

---

* S will give the correct connection even for the phases, because the relative phases of the asymptotic parts are independent of the energy-time factor $e^{-\frac{1}{i\hbar}Et}$, and this factor is expressly eliminated in the interaction representation.

$$\langle \alpha_0'' | S | \alpha_0' \rangle = \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \, U_I(t_0, -\infty) | \alpha_0' \rangle$$

$$= \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \sum_{n=0}^{\infty} \left[-2\pi i \, \delta_-(H_0 - H_0') H_{1s} \right]^n | \alpha_0' \rangle$$

$$= \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \sum_{n=0}^{\infty} \left[-\frac{1}{H_0 - H_0' - i\varepsilon} H_{1s} \right]^n | \alpha_0' \rangle \qquad (10.143)$$

Expansion (10.143) may be co tracted into the following symbolic form :

$$\langle \alpha_0'' | S | \alpha_0' \rangle = \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \left[1 + 2\pi i \, \delta_-(H_0 - H_0') H_{1s} \right]^{-1} | \alpha_0' \rangle$$

$$= \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \left[1 + \frac{1}{H_0 - H_0' - i\varepsilon} H_{1s} \right]^{-1} | \alpha_0' \rangle \qquad (10.144)$$

These are to be compared with the corresponding expansions for the operator $U_I(t_0, -\infty)$:

$$\langle \alpha_0'' | U_I(t_0, -\infty) | \alpha_0' \rangle = \langle \alpha_0'' | \sum_{n=0}^{\infty} \left[-2\pi i \, \delta_-(H_0 - H_0') H_{1s} \right]^n | \alpha_0' \rangle$$

$$= \langle \alpha_0'' | \sum_{n=0}^{\infty} \left[-\frac{1}{H_0 - H_0' - i\varepsilon} H_{1s} \right]^n | \alpha_0' \rangle$$

$$= \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta_-(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \left[1 + 2\pi i \, \delta_-(H_0 - H_0') H_{1s} \right]^{-1} | \alpha_0' \rangle$$

$$= \delta_{\alpha_0'' \alpha_0'} - 2\pi i \, \delta_-(H_0'' - H_0') \langle \alpha_0'' | H_{1s} \left[1 + \frac{1}{H_0 - H_0' - i\varepsilon} H_{1s} \right]^{-1} | \alpha_0' \rangle$$

(10.145)

If the system is in the unperturbed state $|\alpha_0'\rangle$ before scattering, then the probability that it will be found in the unperturbed state $|\alpha_0''\rangle$ after scattering is given by

$$P(\alpha_0'', \alpha_0') = \left| \langle \alpha_0'' | S | \alpha_0' \rangle \right|^2 \qquad (10.146)$$

Since the total probability that the system will be found in some unperturbed state after scattering is unity, we must have the condition

$$1 = \sum_{\alpha_0''}{}' P(\alpha_0'' \leftarrow \alpha_0') = \sum_{\alpha_0''}{}' \langle \alpha_0' | S^* | \alpha_0'' \rangle \langle \alpha_0'' | S | \alpha_0' \rangle$$

$$= \langle \alpha_0' | S^* S | \alpha_0' \rangle \qquad (10.147)$$

Equation (10.147) would certainly be true if the operator S were unitary. When the operator $U_I(t, -\infty)$ is perfectly well behaved -no bound states, etc.- and is hence unitary, it is highly plausible that it remains unitary in the limit $t \to \infty$, and hence that S is unitary. Even when bound states are present and $U_I(t, -\infty)$ is only quasi unitary, the fact that equation (10.132) still holds i implies that (10.147) holds. (Take the limit $t \to \infty$ in (10.132)). We shall later show that even bound states are present the operator S is truly unitary, i.e. that $S^\dagger S = SS^\dagger = 1$.

Let us now therefore investigate the situation that exists when bound states are present. It will be very instructive to carry out our investigation with reference to a very simple system, for which all quantities can be calculated exactly, and which exhibits most of the characteristic features of the theory. Such a system is a one-dimensional particle moving in an attractive delta-function potential, for which the Hamiltonian function is

$$H = \frac{1}{2m} p^2 - \lambda \delta(q) \qquad (10.148)$$

where λ is a positive constant. The unperturbed Hamiltonian function will be taken as that of the free particle,

$$H_0 = \frac{1}{2m} p^2 \qquad (10.149)$$

for which we may use the results of section a) of chapter 9. The

perturbation function is therefore

$$H_1 = -\lambda\,\delta(q) \tag{10.150}$$

Let us work first in the Schrödinger representation and denote by $|H'\rangle$ an eigenvector of the perturbed energy operator H corresponding to the eigenvalue H'. Let us try to take $|H'\rangle$ in such a way that its coordinate representation has the general form

$$\langle q'|H'\rangle = \tfrac{1}{2}\left(1 - \frac{k'q'}{|k'q'|}\right)\left(A\,e^{ik'q'} + B\,e^{-ik'q'}\right) + \tfrac{1}{2}\left(1 + \frac{k'q'}{|k'q'|}\right)C\,e^{ik'q'} \tag{10.151}$$

where A, B, C are certain constants which must be chosen so as to satisfy normalization requirements and the eigenvalue equation $H\,|H'\rangle = H'\,|H'\rangle$. The wave function (10.151) represents a stationary state in which the particle, having momentum $\hbar k'$, comes in toward the origin and has a certain probability of being transmitted through the origin and a certain probability of being reflected back from the origin. The transmission and reflection probability coefficients are given respectively by

$$T = \frac{|C|^2}{|A|^2} \tag{10.152}$$

$$R = \frac{|B|^2}{|A|^2} \tag{10.153}$$

Now, we must have

$$H'\langle q'|H'\rangle = \langle q'|H|H'\rangle = \left[-\frac{\hbar^2}{2m}\frac{d^2}{dq'^2} - \lambda\,\delta(q')\right]\langle q'|H'\rangle \tag{10.154}$$

But

$$\frac{d}{dq'}\langle q'|H'\rangle = -\frac{k'}{|k'|}\,\delta(q')(A+B) + \frac{k'}{|k'|}\,\delta(q')\,C + \tfrac{1}{2}\left(1 - \frac{k'q'}{|k'q'|}\right)\left(ik'A\,e^{ik'q'} - ik'B\,e^{-ik'q'}\right) + \tfrac{1}{2}\left(1 + \frac{k'q'}{|k'q'|}\right)ik'C\,e^{ik'q'} \tag{10.155}$$

If the wave function is to possess no discontinuities we must have

$$C = A + B \tag{10.156}$$

and hence

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{d\rho'^2} - \lambda\delta(\rho')\right]\langle\rho'/H\rangle$$

$$= -\frac{i\hbar^2 k'}{2m}\left[-\frac{k'}{|k'|}\delta(\rho')(A-B) + \frac{k'}{|k'|}\delta(\rho')(A+B)\right]$$

$$+ \frac{\hbar^2 k'^2}{2m}\langle\rho'/H'\rangle - \lambda\delta(\rho')(A+B) \tag{10.157}$$

In order for eqs. (10.154) and (10.157) to be compatible, we must set

$$H' = \frac{\hbar^2 k^2}{2m} \tag{10.158}$$

and

$$-\frac{i\hbar^2|k'|}{m}B - \lambda(A+B) = 0 \tag{10.159}$$

or

$$B = -A\frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}} \tag{10.160}$$

and

$$C = A + B = A\frac{i\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}}. \tag{10.161}$$

The wave function therefore takes the form

$$\langle \rho' | H' \rangle = A \left\{ \left[\tfrac{1}{2}\left(1 - \frac{k\rho'}{|k\rho'|}\right) + \tfrac{1}{2}\left(1 + \frac{k\rho'}{|k\rho'|}\right) \frac{i\frac{\hbar^2 |k'|}{m}}{\lambda + i\frac{\hbar^2 |k'|}{m}} \right] e^{ik\rho'} \right.$$
$$\left. - \tfrac{1}{2}\left(1 - \frac{k\rho'}{|k\rho'|}\right) \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}} e^{-ik\rho'} \right\}. \qquad (10.162)$$

The S-matrix may be read off directly from the asymptotic form of the wave function (10.162)

$$\langle k'' | S | k' \rangle = \frac{i\frac{\hbar^2 |k'|}{m}}{\lambda + i\frac{\hbar^2 |k'|}{m}} \delta_{k''k'} - \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}} \delta_{-k''k'}$$

$$= \delta_{k''k'} - \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}} \left(\delta_{k''k'} + \delta_{-k''k'} \right)$$

$$= \delta_{k''k'} - \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}} \frac{2|k'|}{\delta(k''^2 - k'^2)} \delta(k''^2 - k'^2)$$

$$= \delta_{k''k'} - \delta(H_0'' - H_0') \frac{\hbar^2 |k'| / m}{\delta(k'' - k')} \frac{\lambda}{\lambda + i\hbar^2 k'^2 / m} \qquad (10.163)$$

where we have introduced the eigenvalues of the unperturbed energy operator ;

$$H_0' = \frac{\hbar^2 k'^2}{2m} \quad , \quad H_0'' = \frac{\hbar^2 k''^2}{2m} \, . \qquad (10.164)$$

We see that for these unbound states, $H' = H'_0$, and it is to be anticipated that the state $|H'\rangle$ will be generated from the unperturbed state $|k'\rangle$ where (see (9.20))

$$\langle \rho' | k' \rangle = \sqrt{\frac{\Delta\omega}{L}} \, e^{ik'\rho'} \, . \qquad (10.165)$$

We should therefore relabel the eigenstate $|H'\rangle$ and write it in the form

$$|H'\rangle \equiv |\,\rangle_{k'} \, . \qquad (10.166)$$

The transmission and reflection probability coefficients are now to be given by

$$T = \frac{\left(\frac{\hbar^2 |k'|}{m}\right)^2}{\lambda^2 + \left(\frac{\hbar^2 |k'|}{m}\right)^2}, \quad R = \frac{\lambda^2}{\lambda^2 + \left(\frac{\hbar^2 |k'|}{m}\right)^2} \tag{10.167, 10.168}$$

They evidently satisfy the condition

$$T + R = 1, \tag{10.169}$$

which is another form of the statement of the probability conservation condition (10.147). The unitary of the S-matrix may be verified in this case directly from the form (10.163). We have

$$\begin{aligned}
\langle k''|SS^*|k'\rangle &= \sum_{k'''} \langle k''|S|k'''\rangle \langle k'|S|k'''\rangle^* \\
&= \sum_{k'''} \left(\frac{i\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}} \delta_{k''k'''} - \frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}} \delta_{-k''k'''} \right) \\
&\quad \times \left(\frac{-i\frac{\hbar^2|k'|}{m}}{\lambda - i\frac{\hbar^2|k'|}{m}} \delta_{k'k'''} - \frac{\lambda}{\lambda - i\frac{\hbar^2|k'|}{m}} \delta_{-k'k'''} \right) \\
&= \frac{\left(\frac{\hbar^2|k'|}{m}\right)^2}{\lambda^2 + \left(\frac{\hbar^2|k'|}{m}\right)^2} \delta_{k''k'} - \frac{i\lambda\frac{\hbar^2|k'|}{m}}{\lambda^2 + \left(\frac{\hbar^2|k'|}{m}\right)^2} \delta_{k''-k'} \\
&\quad + \frac{i\lambda\frac{\hbar^2|k'|}{m}}{\lambda^2 + \left(\frac{\hbar^2|k'|}{m}\right)^2} \delta_{-k''k'} + \frac{\lambda^2}{\lambda^2 + \left(\frac{\hbar^2|k'|}{m}\right)^2} \delta_{-k''-k'} \\
&= \delta_{k''k'},
\end{aligned} \tag{10.170}$$

and similarly

$$\langle k''|S^*S|k'\rangle = \delta_{k''k'}. \tag{10.171}$$

The S-matrix is unitary in spite of the fact that when the constant λ is positive there is one bound state. This bound state has the form

$$\langle \rho'|H'\rangle = De^{-a|\rho'|}, \quad a>0 \tag{10.172}$$

where D and a are certain constants. Substituting this form in (10.154) and observing that

$$\frac{d}{d\rho'}\langle \rho'|H'\rangle = -aD\frac{\rho'}{|\rho'|}e^{-a|\rho'|}, \tag{10.173}$$

$$\frac{d^2}{d\rho'^2}\langle \rho'|H'\rangle = -2aD\delta(\rho') + a^2De^{-a|\rho'|}, \tag{10.174}$$

we obtain the equation

$$H'\langle \rho'|H'\rangle = -\frac{\hbar^2a^2}{2m}\langle \rho'|H'\rangle + \frac{\hbar^2a}{m}D\delta(\rho') - \lambda D\delta(\rho'), \tag{10.175}$$

which tells us that we must set

$$a = \frac{\lambda m}{\hbar^2} \tag{10.176}$$

and

$$H' = -\frac{\hbar^2a^2}{2m} = -\frac{\hbar^2}{2m}\frac{\lambda^2m^2}{\hbar^4} = -\frac{\lambda^2m}{2\hbar^2} = E_B \tag{10.177}$$

where we use the symbol E_B to denote the energy of the bound state. The bound state is, of course, non-degenerate, and may therefore be denoted by the single lable E_B:

$$\langle \rho'|E_B\rangle = De^{-\frac{\lambda m}{\hbar^2}|\rho'|} \tag{10.178}$$

We must now investigate the normalization of the stationary state-vectors of the perturbed system. Using (10.62) and (10.66) we have, for k', k" > 0,

$$
\begin{aligned}
{}_{k''}\langle l \rangle_{k'} &= \sum_{\rho'} {}_{k''}\langle \rho' \rangle \langle \rho' |_{k'} \\
&= \frac{|A|^2}{\Delta\omega} \int_{-\infty}^{0} d\rho' \Big\{ e^{i(k'-k'')\rho'} - \frac{\lambda}{\lambda - i\frac{\hbar^2 k''}{m}} e^{i(k'+k'')\rho'} \\
&\quad - \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{-i(k'+k'')\rho'} + \frac{\lambda}{\lambda - i\frac{\hbar^2 k''}{m}} \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{i(k'-k'')\rho'} \Big\} \\
&\quad + \frac{|A|^2}{\Delta\omega} \int_{0}^{\infty} \frac{-i\frac{\hbar^2 k''}{m}}{\lambda - i\frac{\hbar^2 k''}{m}} \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k'}{m}} e^{i(k'-k'')\rho'} d\rho' \\
&= \frac{|A|^2}{\Delta\omega} \pi \Big\{ \delta(k'-k'') + \frac{\lambda^2}{\lambda^2 + \left(\frac{\hbar^2 k''}{m}\right)^2} \delta(k'-k'') + \frac{\left(\frac{\hbar^2 k'}{m}\right)^2}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} \delta(k'-k'') \Big\} \\
&\quad + \frac{|A|^2}{\Delta\omega} i \Big\{ -\frac{1}{k'-k''} + \frac{\lambda}{\lambda - i\frac{\hbar^2 k''}{m}} \frac{1}{k'+k''} - \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} \frac{1}{k'+k''} \\
&\quad + \frac{\lambda^2}{\left(\lambda - i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)} \frac{1}{k'-k''} \\
&\quad + \frac{\frac{\hbar^2 k''}{m} \frac{\hbar^2 k'}{m}}{\left(\lambda - i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)} \frac{1}{k'-k''}
\end{aligned}
$$

$$= \frac{|A|^2}{\Delta\omega} 2\pi\delta(k'-k'') + \frac{|A|^2}{\Delta\omega} \frac{i}{\left(\lambda - i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)(k'-$$

$$\times \left\{ -\left(\lambda - i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right) + i\lambda\frac{\hbar^2}{m}(k'-k'') + \lambda^2 + \frac{\hbar^2 k''}{m} \right.$$

$$= \frac{|A|^2}{\Delta\omega} 2\pi\delta(k'-k'') = \frac{|A|^2}{\Delta\omega} L\,\delta_{k'k''}\,, \qquad (10.$$

and for $k' > 0$, $k'' < 0$,

$$\left\langle \right|_{k''} \left| \right\rangle_{k'} = \frac{|A|^2}{\Delta\omega} \int_{-\infty}^{0} \left\{ \frac{i\frac{\hbar^2 k''}{m}}{\lambda + i\frac{\hbar^2 k'}{m}} e^{i(k-k'')\rho'} - \frac{i\frac{\hbar^2 k''}{m}}{\lambda + i\frac{\hbar^2 k''}{m}} \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{-i(k'+k'')\rho'} \right\} d$$

$$+ \frac{|A|^2}{\Delta\omega} \int_{0}^{\infty} \left\{ \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k'}{m}} e^{i(k'-k'')\rho'} - \frac{\lambda}{\lambda + i\frac{\hbar^2 k''}{m}} \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k'}{m}} e^{i(k'+k'')\rho'} \right\} d$$

$$= \frac{|A|^2}{\Delta\omega} \pi \left\{ \frac{i\frac{\hbar^2 k'}{m}\lambda}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} \delta(k'-k'') - \frac{\lambda i\frac{\hbar^2 k'}{m}}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} \delta(k'+k'') \right.$$

$$+ \frac{|A|^2}{\Delta\omega} i \left\{ -\frac{i\frac{\hbar^2 k''}{m}}{\lambda + i\frac{\hbar^2 k''}{m}} \frac{1}{k'-k''} - \frac{i\frac{\hbar^2 k''}{m}\lambda}{\left(\lambda + i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)} \frac{1}{k'+k''} \right.$$

$$\left. + \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k''}{m}} \frac{1}{k'-k''} - \frac{\lambda i\frac{\hbar^2 k'}{m}}{\left(\lambda + i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)} \frac{1}{k'+k''} \right\}$$

$$= \frac{|A|^2}{\Delta\omega} \frac{i}{\left(\lambda + i\frac{\hbar^2 k''}{m}\right)\left(\lambda + i\frac{\hbar^2 k'}{m}\right)(k'-k'')}$$

$$\times \left\{ -i\frac{\hbar^2 k''}{m}\left(\lambda + i\frac{\hbar^2 k'}{m}\right) + i\frac{\hbar^2 k'}{m}\left(\lambda + i\frac{\hbar^2 k''}{m}\right) - i\frac{\hbar^2}{m}\lambda(k'-k'') \right\}$$

$$= 0.$$

(10.180)

From symmetry it is evident that we also have

$$\langle k'' | \rangle_{k'} = \frac{|A|^2}{\Delta\omega} L\,\delta_{k'k''} \quad \text{for } k', k'' < 0 \qquad (10.181)$$

and

$$\langle k'' | \rangle_{k'} = 0 \quad \text{for } k' < 0,\ k'' > 0. \qquad (10.182)$$

For normalization it is evident that we may choose the constant A to be

$$A = \sqrt{\frac{\Delta\omega}{L}}. \qquad (10.183)$$

Hence

$$\langle \rho' | \rangle_{k'} = \sqrt{\frac{\Delta\omega}{L}} \left\{ \left[\tfrac{1}{2}\left(1 - \frac{k'\rho'}{|k'\rho'|}\right) + \tfrac{1}{2}\left(1 + \frac{k'\rho'}{|k'\rho'|}\right) \frac{i\frac{\hbar^2 |k'|}{m}}{\lambda + i\frac{\hbar^2 |k'|}{m}} \right] e^{ik'\rho'} \right.$$

$$\left. - \tfrac{1}{2}\left(1 - \frac{k'\rho'}{|k'\rho'|}\right) \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}}\, e^{-ik'\rho'} \right\}. \qquad (10.184)$$

For the bound state we have

$$\langle E_B | E_B \rangle = \sum_{\rho'} \langle E_B | \rho' \rangle \langle \rho' | E_B \rangle$$

$$= \frac{|D|^2}{\Delta\omega} \int_{-\infty}^{\infty} e^{-\frac{2\lambda m}{\hbar^2}|\rho'|} d\rho' = \frac{|D|^2}{\Delta\omega} 2\left(-\frac{\hbar^2}{2\lambda m}\right)\left[e^{-\frac{2\lambda m}{\hbar^2}\rho'}\right]_0^{\infty} \quad (10.185)$$

$$= \frac{|D|^2}{\Delta\omega} \frac{\hbar^2}{\lambda m}.$$

For normalization it is evident that we may choose

$$D = \sqrt{\frac{\lambda m}{\hbar^2} \Delta\omega}, \quad (10.186)$$

so that

$$\langle \rho' | E_B \rangle = \sqrt{\frac{\lambda m}{\hbar^2} \Delta\omega}\; e^{-\frac{\lambda m}{\hbar^2}|\rho'|} \quad (10.187)$$

One may readily show that the bound state vector is orthogonal to the unbound state vectors : for $k' > 0$

$$\langle E_B | \rangle_{k'}$$

$$= \sqrt{\frac{\lambda m}{\hbar^2 L}} \int_{-\infty}^{\infty} e^{-\frac{\lambda m}{\hbar^2}|\rho'|} \left\{ \left[\frac{1}{2}\left(1 - \frac{\rho'}{|\rho'|}\right) + \frac{1}{2}\left(1 + \frac{\rho'}{|\rho'|}\right) \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k'}{m}}\right] e^{ik'\rho'} \right.$$

$$\left. - \frac{1}{2}\left(1 - \frac{\rho'}{|\rho'|}\right) \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{-ik'\rho'} \right\} d\rho'$$

$$= \sqrt{\frac{\lambda m}{\hbar^2 L}} \left\{ \frac{1}{\frac{\lambda m}{\hbar^2} + ik'} \left[e^{\left(\frac{\lambda m}{\hbar^2} + ik'\right)\rho'} \right]_{-\infty}^{0} - \frac{1}{\frac{\lambda m}{\hbar^2} - ik'} \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} \left[e^{\left(\frac{\lambda m}{\hbar^2} - ik'\right)\rho'} \right]_{-\infty}^{0} \right.$$

$$\left. + \frac{1}{-\frac{\lambda m}{\hbar^2} + ik'} \frac{i\frac{\hbar^2 k'}{m}}{\lambda + i\frac{\hbar^2 k'}{m}} \left[e^{\left(-\frac{\lambda m}{\hbar^2} - ik'\right)\rho'} \right]_{-\infty}^{0} \right\} \quad (10.188)$$

$$= \sqrt{\frac{\lambda m}{\hbar^2 L}} \frac{\hbar^2}{m} \left\{ \frac{1}{\lambda + i\hbar^2 k'/m} - \frac{\lambda}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} + \frac{i\frac{\hbar^2 k'}{m}}{\lambda^2 + (\hbar^2 k'/m)^2} \right\} = 0.$$

and, by symmetry, also $\langle E_B | \rangle_{k'} = 0$ for $k' < 0$.

For later work it will be useful to have the momentum-space representations of the eigenvectors of H. Using (9.20) and (10.184), we have

$$\langle k'' | \rangle_{k'} = \sum_{\rho'} \langle k'' | \rho' \rangle \langle \rho' | \rangle_{k'}$$

$$= \frac{1}{L} \int_{-\infty}^{\infty} e^{-ik''\rho'} \left\{ \left[\tfrac{1}{2}\left(1 - \frac{k'\rho'}{|k'\rho'|}\right) + \tfrac{1}{2}\left(1 + \frac{k'\rho'}{|k'\rho'|}\right) \frac{\frac{\hbar^2 |k'|}{m}}{\lambda + i\frac{\hbar^2 |k'|}{m}} \right] e^{ik'\rho'} \right.$$

$$\left. - \frac{1}{2}\left(1 - \frac{k'\rho'}{|k'\rho'|}\right) \frac{\lambda}{\lambda + i\frac{\hbar^2 |k'|}{m}} e^{-ik'\rho'} \right\} d\rho'$$

$$= \frac{1}{L} \left\{ -i \frac{k'}{|k'|} \frac{1}{k' - k'' - i\epsilon \frac{k'}{|k'|}} + i \frac{k'}{|k'|} \frac{i\hbar^2 |k'|/m}{\lambda + i\hbar^2 |k'|/m} \frac{1}{k' - k'' + i\epsilon \frac{k'}{|k'|}} \right.$$

$$\left. - i \frac{k'}{|k'|} \frac{\lambda}{\lambda + i\hbar^2 |k'|/m} \frac{1}{k' - k'' + i\epsilon \frac{k'}{|k'|}} \right\}, \qquad (10.189)$$

$$\langle k'' | E_B \rangle = \sqrt{\frac{\lambda m}{\hbar^2 L}} \int_{-\infty}^{\infty} e^{-ik''\rho'} e^{-\lambda \frac{m}{\hbar^2}|\rho'|} d\rho'$$

$$= \sqrt{\frac{\lambda m}{\hbar^2 L}} \left\{ \frac{1}{\frac{\lambda m}{\hbar^2} - ik''} + \frac{1}{\frac{\lambda m}{\hbar^2} + ik''} \right\}$$

$$= \sqrt{\frac{\lambda m}{\hbar^2 L}} \, 2 \frac{\hbar^2}{m} \lambda \frac{1}{\lambda^2 + \left(\frac{\hbar^2 k''}{m}\right)^2}$$

$$= \sqrt{\frac{\hbar^2}{\lambda m L}} \frac{2\lambda^{\frac{3}{2}}}{\lambda^2 + \left(\frac{\hbar^2 k''}{m}\right)^2} \qquad (10.190)$$

One may also check the completeness of the vectors

$|\rangle_{\mu_0 |k'|}$, $|E_B\rangle$ in a straightforward manner. For ρ', $\rho'' > 0$ we have

$$\sum_{k'} \langle \rho'' | \rangle_{k'} \langle k' | \rho' \rangle = \frac{L}{2\pi} \int_{-\infty}^{\infty} \langle \rho'' | \rangle_{k'} \langle k' | \rho' \rangle \, dk'$$

$$= \frac{\Delta\omega}{2\pi} \int_{-\infty}^{0} \left\{ e^{-ik'(\rho'-\rho'')} - \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{ik'(\rho'+\rho'')} - \frac{\lambda}{\lambda - i\frac{\hbar^2 k'}{m}} e^{ik'(\rho'+\rho'')} + \frac{\lambda^2}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} e^{ik'(\rho'+\rho'')} \right\} dk'$$

$$+ \frac{\Delta\omega}{2\pi} \int_{0}^{\infty} \frac{\left(\frac{\hbar^2 k'}{m}\right)^2}{\lambda^2 + \left(\frac{\hbar^2 k'}{m}\right)^2} e^{-i(\rho'-\rho'')k'} dk'$$

$$= \frac{\Delta\omega}{2\pi} \int_{-\infty}^{\infty} \left\{ e^{-ik'(\rho'-\rho'')} - \frac{\lambda}{\lambda + i\frac{\hbar^2 k'}{m}} e^{ik'(\rho'+\rho'')} \right\} dk'$$

$$= \Delta\omega \, \delta(\rho'-\rho'') + \frac{\Delta\omega}{2\pi} i \frac{\lambda m}{\hbar^2} \int_{-\infty}^{\infty} \frac{1}{k' - i\lambda m/\hbar^2} e^{ik'(\rho'+\rho'')} dk' \quad (*)$$

$$= \delta_{\rho'\rho''} - \Delta\omega \frac{\lambda m}{\hbar^2} e^{-\lambda m/\hbar^2 (\rho'+\rho'')} \qquad (10.19)$$

For $\rho' > 0$, $\rho'' < 0$ we have

$$\sum_{k'} \langle \rho'' | \rangle_{k'} \langle k' | \rho' \rangle$$

---

* When λ is negative the pole of the integrand lies below the real axis and the integral vanishes. There are no bound states, and the unbound states by themselves form a complete set.

$$= \frac{\Delta\omega}{2\pi}\int_{-\infty}^{0}\left\{\frac{-i\frac{\hbar^2k'}{m}}{\lambda-i\frac{\hbar^2k'}{m}}\right. e^{-ik'(\rho'-\rho'')}\frac{-i\frac{\hbar^2k'}{m}}{\lambda-i\frac{\hbar^2k'}{m}}\frac{\lambda}{\lambda+i\frac{\hbar^2k'}{m}}\, e^{ik'(\rho'-\rho'')}$$

$$+\frac{\Delta\omega}{2\pi}\int_{0}^{\infty}\left\{\frac{-i\frac{\hbar^2k'}{m}}{\lambda-i\frac{\hbar^2k'}{m}}\right. e^{-ik'(\rho'-\rho'')}\frac{-i\frac{\hbar^2k'}{m}}{\lambda-i\frac{\hbar^2k'}{m}}\frac{\lambda}{\lambda+i\frac{\hbar^2k'}{m}}\, e^{-ik'(\rho'-\rho'')}$$

$$= \frac{\Delta\omega}{2\pi}\int_{-\infty}^{\infty}\left\{\frac{i\hbar^2k'/m}{\lambda+i\hbar^2k'/m}\right. e^{ik'(\rho'-\rho'')}dk' = \frac{\Delta\omega}{2\pi}\int_{-\infty}^{\infty}\frac{k'}{k'-i\frac{\lambda m}{\hbar^2}}\, e^{ik'(\rho'-\rho'')}$$

$$= -\Delta\omega\,\frac{\lambda m}{\hbar^2}\, e^{-\frac{\lambda m}{\hbar^2}(\rho'-\rho'')}, \qquad (10.19\ldots)$$

and, for symmetry, we see that in general

$$\sum_{k'}\langle\rho''|k'\rangle\langle k'|\rho'\rangle = \delta_{\rho\rho'} - \Delta\omega\,\frac{\lambda m}{\hbar^2}\, e^{-\frac{\lambda m}{\hbar^2}(\rho''-\rho')} \qquad (10.19\ldots)$$

Hence

$$\sum_{k'}\langle\rho''|k'\rangle\langle k'|\rho'\rangle + \lambda\langle\rho''|G_0|\rho'\rangle\langle E_0(\rho')\rangle = \delta_{\rho\rho'}, \qquad (10.194)$$

and the completeness is thus established.

In order to calculate the λ expansions for the operators $U_I(t_0,-\infty)$, S, etc., it will be necessary to have the momentum-space matrix elements of the interaction operator H_{I_S}. We have

$$\langle k''|H_{I_S}|k'\rangle = -\lambda\langle k''|\delta(\rho)|k'\rangle$$

$$= -\lambda\sum_{\rho}\langle k''|\rho\rangle\langle\rho|\delta(\rho)|\rho'\rangle\langle\rho'|k'\rangle$$

$$= -\frac{\lambda}{L}\int_{-\infty}^{\infty}\delta(\rho)\, e^{-i(k'-k'')\rho'}\,d\rho' = -\frac{\lambda}{L}. \qquad (10.195)$$

That is, all the matrix elements are equal. This fact will greatly simplify the calculation of the various expansions.

Using (10.143) we have

$$\langle k''|S|k'\rangle = \delta_{k''k'} - 2\pi i\,\delta(H_0''-H_0')\langle k''|H_{1s}\sum_{n=0}^{\infty}\left[-\frac{1}{H_0-H_0'-i\varepsilon}H_{1s}\right]^n$$

$$= \delta_{k''k'} + 2\pi i\,\delta(H_0''-H_0')\frac{\lambda}{L}\sum_{n=0}^{\infty}\left\{\left(\frac{\lambda}{L}\right)^n\sum_{k^{(1)}\cdots k^{(n)}}\frac{1}{H_0^{(1)}-H_0'-i\varepsilon}\cdots\frac{1}{H_0^{(n)}-H_0'-i\varepsilon}\right\}$$

$$= \delta_{k''k'} + i\,\delta(H_0''-H_0')\frac{\lambda}{\delta(k'-k')}\sum_{n=0}^{\infty}\left\{\left(\frac{\lambda}{2\pi}\right)^n\left(\frac{2m}{\hbar^2}\right)^n \int_{-\infty}^{\infty}dk^{(1)}\cdots\int_{-\infty}^{\infty}dk^{(n)}\frac{1}{k^{(1)2}-k'^2-i\underline{\varepsilon}}\cdots\frac{1}{k^{(n)2}-k'^2-i\underline{\varepsilon}}\right\} \quad (10.19)$$

$$\left[\underline{\varepsilon} = \frac{2m}{\hbar^2}\varepsilon\right]$$

But

$$\int_{-\infty}^{\infty}\frac{dk}{k^2-k'^2-i\varepsilon} = \int_{-\infty}^{\infty}\frac{dk}{(k-\sqrt{k'^2+i\varepsilon})(k+\sqrt{k'^2+i\varepsilon})} = \frac{\pi i}{\sqrt{k'^2+i\varepsilon}} \quad (10.19')$$

Hence, in the limit $\varepsilon \to 0$, we have

$$\langle k''|S|k'\rangle = \delta_{k''k'} + i\,\delta(H_0''-H_0')\frac{\lambda}{\delta(k'-k')}\sum_{n=0}^{\infty}\left(i\frac{\lambda m}{\hbar^2}\right)^n\frac{1}{|k'|^n}, \quad (10.19'')$$

The series expansion clearly converges only for values of k' which satisfy

$$|k'| > \frac{\lambda m}{\hbar^2} \quad (10.1\ldots)$$

This corresponds to values of energy satisfying

$$H_0' > \frac{\hbar^2}{2m}\frac{\lambda^2 m^2}{\hbar^4} = \frac{\lambda^2 m}{2\hbar^2} = |E_B|. \quad (10.20)$$

That is, the series expansion for the S-matrix will converge only [illegible]

incoming particle has an energy greater than the absolute value of the energy of the bound state. When condition (10.200) is satisfied (10.198) may be written in the form

$$\langle k''|S|k'\rangle = \delta_{k''k'} + i\delta(H''_0 - H'_0)\frac{\lambda}{\delta(k'-k)}\,\frac{1}{\lambda - i\frac{\hbar^2|k'|}{m}}$$

$$= \delta_{k''k'} + i\delta(H''_0 - H'_0)\frac{1}{\delta(k'-k)}\,\frac{i\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}} \qquad (10.20$$

which is identical with (10.163).

The matrix elements of the operator $U_I(t_0, -\infty)$ may be read off directly from equation (10.189) if we write it in the form

$$\langle k''|U_I(t_0,-\infty)|k'\rangle = \langle k''|\lambda_1$$

$$= \frac{1}{2\delta(k'-k)}\left\{\delta(k'-k'') + \frac{i\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}}\,\delta(k'-k'') - \frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}}\,\delta(k'-k'')\right.$$

$$\left. + \frac{i}{L}\frac{k'}{|k'|}\left\{-\frac{1}{k'-k''} + i\frac{\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}}\,\frac{1}{k'-k''} - \frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}}\,\frac{1}{k'+k''}\right.\right.$$

$$= \frac{1}{2}\left[1 + \frac{i\frac{\hbar^2|k'|}{m}}{\lambda + i\frac{\hbar^2|k'|}{m}}\right]\delta_{k''k'} - \frac{1}{2}\,\frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}}\,\delta_{-k''k'}$$

$$- \frac{i}{L}\frac{k'}{|k'|}\,\frac{\lambda}{\lambda - i\frac{\hbar^2|k'|}{m}}\left[\frac{1}{k'-k''} - \frac{1}{k'+k''}\right]$$

$$= \delta_{k''k'} - \frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}}\left[\frac{1}{2}\delta_{k''k'} - \frac{1}{2}\delta_{-k''k'} + \frac{i}{L}\frac{k'}{|k'|}\left(\frac{1}{k'-k''} + \frac{1}{k'+k''}\right)\right]$$

$$= \delta_{k''k'} - \frac{2|k'|}{\delta(k''-k')}\,\frac{\lambda}{\lambda + i\frac{\hbar^2|k'|}{m}}\left[\frac{1}{2}\delta(k''^2 - k'^2) - \frac{i}{2\pi}\,\frac{1}{k''^2 - k'^2}\right]$$

$$= \delta_{k''k_0} - \delta(H_0'' - H_0') \frac{\lambda^n \Pi^n}{\Gamma(k'-n')} \frac{\partial}{\partial \alpha^n} \frac{1}{\alpha + i(k'-n')} \qquad (10.202)$$

This is to be compared with (10.163). It is obvious that if we use (10.145) to compute the mayrix elements of $U_I(t_0, -\infty)$, the computations are exactly the same as those for the S-matrix and we obtain the binomial expansion of (10.202). Again the expansion converges only if the condition (10.200) is satisfied.

It should be now observed, from the second form for $\langle k'' | U_I(t_0, -\infty) | k' \rangle$ in (10.202), that when k' = 0, $\langle k'' | U_I(t_0, -\infty) | k' \rangle$ vanishes for all k". This implies (cf.(10.66)

$$U_I(t_0, -\infty) | c \rangle = 0 \qquad (10.203)$$

which means that the operator $U_I(t_0, -\infty)$ transforms the unperturbed state 10) into no state at all. This may also be seen from the wave function (10.184), which vanishes for k' = 0. That is, there is no perturbed stationary corresponding to the value k' = 0

What happens to this state ? Naturally, we have a strong suspicion that it gets "peeled off of the bottom of the deck" of unperturbed continuum states and "sucked in" by the potential $-\lambda\delta(\xi)$ to form the single bound perturbed state. This one particular state therefore suffers an energy shift, and obviously some catastrophe such as (10.203) must occur.

We shall verify this suspicion by discussing the whole situation step by step in considerable detail. We shall first take a look at the unitarity of the operator $U_I(t_0, -\infty)$. Consider the perturbed state vectors $| \rangle_{k'}$. We have already proved that these vectors are orthonormal. Hence we may write

$$\delta_{k'k'} = {}_{k'}\langle | \rangle_{k'} = \langle k'' | U_I^*(t_0, -\infty) U_I(t_0, -\infty) | k' \rangle \qquad (10.204)$$

Since this equation holds for all k', k", and since the unperturbed vectors $| k' \rangle$ form a complete set, we have verified that equation

(10.132) holds in the present case. The equation $U_I(t_0, -\infty)$ $U^*(t_0, -\infty) = 1$, on the other hand, cannot hold. For if it did, the vectors $|k'\rangle$ would form a complete set, and we have seen that they form a complete set only if the bound state vector $|E_B\rangle$ is added to them. That is, we have

$$1 \neq \sum_{k'} |k'\rangle_{k'} \; {}_{k'}\langle 1|$$

$$= \sum_{k'} U_I(t_0, -\infty)|k'\rangle\langle k'|U_I^*(t_0, -\infty) = U_I(t_0, -\infty)U_I^*(t_0, -\infty) \quad (10.205)$$

The discussion here is quite general and holds for all systems which possess bound as well as continuum stationary states.

The next point to observe is that equation (10.203) is really not correct. For if the operator $U_I(t_0, -\infty)$ is defined by the expansion (10.48), then, in the limit $\epsilon \to +0$, the vector $U_I(t_0, -\infty)|0\rangle$ is not null, but merely undetermined, or non-convergent. This suggests that we can actually obtain the bound state by taking a ratio of the form (10.119). That is, we suspect that $|E_B\rangle$ can be given by an expression of the form

$$C|E_B\rangle = \frac{U_I(t_0, -\infty)|0\rangle}{\langle 0|U_I(t_0, -\infty)|0\rangle}, \quad (10.206)$$

where C is a certain constant which will account for the fact that the expression on the right is not, in general, normalized. We shall now verify equation (10.206) directly. Using (10.48), (10.195) and (10.197), we have

$$\langle k'|U_I(t_0, -\infty)|0\rangle =$$

$$\delta_{k'0} + \sum_{n=1}^{\infty} \frac{(-1)^n}{H_0' - i\eta\epsilon} \langle k'|H_1' \frac{1}{H_0 - i(n-1)\epsilon} \cdots$$

$$\cdots \frac{1}{H_0 - i\epsilon} H_1'|0\rangle$$

$$= \delta_{k'0} + \sum_{n=1}^{\infty} \left(\frac{\lambda}{L}\right)^n \frac{1}{+k_0' - in\epsilon} \sum_{k^{(1)} \cdots k^{(n-1)}} \frac{1}{+k_0^{(n-1)} - i(n-1)\epsilon} \cdots \frac{1}{+k_0^{(1)} - i\epsilon}$$

$$= \delta_{k'0} + \sum_{n=1}^{\infty} \frac{2\pi}{L} \left(\frac{\lambda}{2\pi}\right)^n \left(\frac{2m}{\hbar^2}\right)^n \frac{1}{k'^2 - in\epsilon} \int_{-\infty}^{\infty} dk^{(1)} \cdots \int_{-\infty}^{\infty} dk^{(n-1)} \frac{1}{k^{(n-1)^2} - (n-1)\epsilon} \cdots \frac{1}{k^{(1)^2} - i\epsilon}$$

$$= \delta_{k'0} + \frac{2\lambda m}{L\hbar^2} \sum_{n=0}^{\infty} \left(i\frac{\lambda m}{\hbar^2}\right)^n \frac{1}{k'^2 - i(n+1)\epsilon} \frac{1}{\sqrt{in\epsilon}} \cdots \frac{1}{\sqrt{L\epsilon}}$$

$$= \delta_{k'0} + \frac{2\hbar^2}{L\lambda m} F(K, \alpha) \qquad (10.207)$$

where

$$F(K,\alpha) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{1}{K^2 + (n+1)\alpha^2} \frac{1}{\alpha^n} \qquad (10.208)$$

with

$$K = \frac{\hbar^2 k'}{\lambda m} \qquad (10.209)$$

and

$$\alpha = \sqrt{\frac{\epsilon}{i}} \frac{\hbar^2}{\lambda m} = \sqrt{\frac{2\hbar^2}{\lambda^2 m} \frac{\epsilon}{i}} = \sqrt{\frac{i\epsilon}{E_B}} \,. \qquad (10.210)$$

The limit $\epsilon \to 0$ may now be expressed in the form $\alpha \to 0$.

We shall now obtain an symptotic form for the series (10.208). First consider the series

$$G(\alpha) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{1}{\alpha^n}$$

$$= \sum_{m=0}^{\infty} \frac{1}{\sqrt{(2m)!}} \frac{1}{\alpha^{2m}} + \sum_{m=0}^{\infty} \frac{1}{\sqrt{(2m+1)!}} \frac{1}{\alpha^{2m+1}} \qquad (10.211)$$

If we use the Stirling asymptotic formula,

$$n! \sim \sqrt{2\pi}\; n^{n+\frac{1}{2}} e^{-n}, \tag{10.212}$$

we may write

$$\sqrt{(2m)!} \sim (2\pi)^{\frac{1}{4}} (2m)^{m+\frac{1}{4}} e^{-m}$$

$$\sim \frac{2^m}{(m\pi)^{\frac{1}{4}}} m!, \tag{10.213}$$

and,

$$\sqrt{(2m+1)!} \sim (2\pi)^{\frac{1}{4}} (2m+1)^{m+\frac{3}{4}} e^{-m-\frac{1}{2}}$$

$$= (2\pi)^{\frac{1}{4}} (2m)^{m+\frac{3}{4}} \left(1+\frac{1}{2m}\right)^{m+\frac{3}{4}} e^{-m-\frac{1}{2}}$$

$$\sim (2\pi)^{\frac{1}{4}} (2m)^{m+\frac{3}{4}} e^{-m}$$

$$\sim 2^{\frac{1}{2}} \left(\frac{m}{\pi}\right)^{\frac{1}{4}} 2^m m! \tag{10.214}$$

Moroover, if α is real, positive and very small, the terms of the series (10.211) do not begin to take on significant values until the following condition is satisfied :

$$0 = \frac{\partial}{\partial n} \log\left(\sqrt{n!}\,\alpha^n\right)$$

$$\sim \frac{\partial}{\partial n}\left[\tfrac{1}{2}\left(n+\tfrac{1}{2}\right)\log n - \tfrac{1}{2}n + n\log\alpha\right]$$

$$= \tfrac{1}{2}\log n + \tfrac{1}{2}\frac{n+\frac{1}{2}}{n} - \tfrac{1}{2} + \log\alpha \sim \log\alpha\sqrt{n} \tag{10.215}$$

or

$$n \sim 2m \sim \frac{1}{\alpha^2} \tag{10.216}$$

Hence, when α is real and positive, we may write

$$G(\alpha) \sim \sum_{m=0}^{\infty} \frac{(m\pi)^{\frac{1}{4}}}{m!} \frac{1}{(2\alpha^2)^m} + \sum_{m=0}^{\infty} 2^{-\frac{1}{2}} \left(\frac{\pi}{m}\right)^{\frac{1}{4}} \frac{1}{\alpha} \frac{1}{m!} \frac{1}{(2\alpha^2)^m} \quad (10.217)$$

For very small α we may replace the factors $m^{\frac{1}{4}}$, $m^{-\frac{1}{4}}$ in the series above by the values which these factors assume for the largest term, as given by (10.216). Hence

$$G(\alpha) \sim 2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \frac{1}{\sqrt{\alpha}} \sum_{m=0}^{\infty} \frac{1}{m!} \frac{1}{(2\alpha^2)^m}$$

$$= 2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \frac{1}{\sqrt{\alpha}} e^{\frac{1}{2\alpha^2}} \quad (10.218)$$

Although we have obtained the asymptotic form (10.218) assuming α real and positive, it holds, by analytic continuation, for complex values of α.

We may now write

$$\alpha \frac{d}{d\alpha}\left(\frac{1}{\alpha} G(\alpha)\right) \sim 2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \alpha \left(-\frac{3}{2\alpha^{5/2}} - \frac{1}{\alpha^{9/2}}\right) e^{\frac{1}{2\alpha^2}}$$

$$\sim -2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \frac{1}{\alpha^{\frac{7}{2}}} e^{\frac{1}{2\alpha^2}}, \quad (10.219)$$

$$\left(\alpha \frac{d}{d\alpha}\right)^2 \left(\frac{1}{\alpha} G(\alpha)\right) \sim -2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \alpha \left(-\frac{7}{2\alpha^{9/2}} - \frac{1}{\alpha^{13/2}}\right) e^{\frac{1}{2\alpha^2}}$$

$$\sim 2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \frac{1}{\alpha^{\frac{11}{2}}} e^{\frac{1}{2\alpha^2}} \quad (10.220)$$

. . . . . .

$$\left(\alpha \frac{d}{d\alpha}\right)^m \left(\frac{1}{\alpha} G(\alpha)\right) \sim 2\left(\frac{\pi}{2}\right)^{\frac{1}{4}} \frac{(-1)^m}{\alpha^{2m+\frac{3}{2}}} e^{\frac{1}{2\alpha^2}} \quad (10.221)$$

But

$$\frac{1}{\alpha} G(\alpha) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{1}{\alpha^{n+1}} , \tag{10.222}$$

$$\alpha \frac{d}{d\alpha} \left(\frac{1}{\alpha} G(\alpha) \right) = - \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{n+1}{\alpha^{n+1}} , \tag{10.223}$$

. . . . . .

$$\left(\alpha \frac{d}{d\alpha} \right)^m \left(\frac{1}{\alpha} G(\alpha) \right) = (-1)^m \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{(n+1)^m}{\alpha^{n+1}} . \tag{10.224}$$

The function $F(K,\alpha)$ may therefore be written in the asymptotic form

$$\begin{aligned} F(K,\alpha) &= \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \frac{1}{K^2} \left\{ 1 - (n+1)\frac{\alpha^2}{K^2} + (n+1)^2 \frac{\alpha^4}{K^4} - \dots \right\} \frac{1}{\alpha^n} \\ &= \frac{1}{K^2} \sum_{m=0}^{\infty} \alpha \left(\frac{\alpha^2}{K^2} \right)^m \left(\alpha \frac{d}{d\alpha} \right)^m \left(\frac{1}{\alpha} G(\alpha) \right) \\ &\sim \left(\frac{\pi}{2} \right)^{\frac{1}{4}} \frac{2}{K^2} \sum_{m=0}^{\infty} (-1)^m \frac{1}{\alpha^{2m+\frac{1}{2}}} \left(\frac{\alpha^2}{K^2} \right)^m e^{\frac{1}{2\alpha^2}} \\ &= 2 \left(\frac{\pi}{2} \right)^{\frac{1}{4}} \frac{\frac{1}{\sqrt{\alpha}} e^{\frac{1}{2\alpha^2}}}{K^2+1} \end{aligned} \tag{10.225}$$

We thus finally have

$$\begin{aligned} \langle k' | U_I(t_0,\infty) | 0 \rangle &\sim \left(\frac{\pi}{2} \right)^{\frac{1}{4}} \frac{4\hbar^2}{L\lambda m} \frac{1}{\sqrt{\alpha}} \frac{e^{\frac{1}{2\alpha^2}}}{K^2+1} \\ &= 2 \left(\frac{\pi}{2} \right)^{\frac{1}{4}} \frac{\hbar^2}{\lambda m L} \frac{2\rangle^2}{\lambda^2 + \left(\frac{\hbar^2 k'}{m} \right)^2} \left(\frac{E_B}{iE} \right)^{\frac{1}{4}} e^{\frac{1}{2}\frac{E_B}{iE}} , \end{aligned} \tag{10.226}$$

and in particular

$$\langle 0|U_I(t_0,-\infty)|0\rangle \sim 4\left(\frac{\pi}{2}\right)^{\frac{1}{4}}\frac{\hbar^2}{\lambda m L}\left(\frac{E_B}{i\epsilon}\right)^{\frac{1}{4}} e^{\frac{1}{2}\frac{E_B}{i\epsilon}} \qquad (10.227)$$

Hence

$$\frac{\langle \hbar_2'|U_I(t_0,-\infty)|0\rangle}{\langle 0|U_I(t_0,-\infty)|0\rangle} = \frac{\lambda^2}{\lambda^2+\left(\frac{\hbar^2\hbar_2'}{m}\right)^2} \qquad (10.228)$$

Remembering that E_B is proportional to λ^2 (see (10.177)), we have, on applying equation (10.121) to (10.227),

$$\Delta E = i\epsilon\lambda\frac{\partial}{\partial\lambda}\log\langle 0|\rangle_0 = i\epsilon\lambda\frac{\partial}{\partial\lambda}\log\langle 0|U_I(t_0,-\infty)|0\rangle$$
$$= i\epsilon\lambda\frac{\partial}{\partial\lambda}\left[\frac{1}{4}\log\frac{E_B}{i\epsilon}+\frac{1}{2}\frac{E_B}{i\epsilon}\right] = \frac{1}{2}\lambda\frac{\partial}{\partial\lambda}E_B = E_B$$

which checks.

Comparison of (10.206) and (10.228) with (10.190) indicates that the normalizing constant C has the value

$$C = \frac{1}{2}\sqrt{\frac{\lambda m L}{\hbar^2}} \qquad (10.229)$$

It will also be observed that if we make the formal identification

$$2\left(\frac{\pi}{2}\right)^{\frac{1}{4}}\left(\frac{E_B}{i\epsilon}\right)^{\frac{1}{4}} e^{\frac{1}{2}\frac{E_B}{i\epsilon}} = \sqrt{\frac{\lambda m L}{\hbar^2}} \qquad (10.230)$$

we may write immdediately

$$U_I(t_0,-\infty)|0\rangle = |E_B\rangle \qquad (10.231)$$

in place of (10.203). In this way we can, if we wish, force the unitary property back into $U_I(t_0, -\infty)$.

The results obtained for the one-dimensional delta function potential are typical of the situation which occurs in the general case when bound states are present. We can in fact characterize the general case qualitatively as follows. Suppose the unperturbed system has no bound stationary states, and suppose the perturbed system has just one bound state. Let the parameter λ, to which the perturbation function H_1 is imagined as being proportional, be gradually "switched off". The level E_B of the bound state will rise during this process. At a certain point the bound state will no longer exist, it having the value of the parameter λ when this event occurs, and let E_0 be the corresponding value of the bound state energy level, i.e. $E_0 = \lim_{\lambda \to \lambda_0} E_B$. In the case of the one-dimensional delta function considered above $\lambda_0 = 0$ and $E_0 = 0$. One can, if one wishes, assume in the general case that the zero point of energy has been chosen in such a way that $E_0 = 0$ The energies of all the continuum states will then lie above zero. One is not, on the other hand, at liberty to choose $\lambda_0 = 0$ in general. For, it may frequently happen (although the opposite also frequently happens) that as the perturbation is "switched on", nothing happens until the parameter λ reaches a certain point λ_0, when suddenly a bound state is born. It just happens that in the case of all one-dimensional potentials this event occurs instantaneously at $\lambda = 0$

By analogy with the example considered above, one might suppose that the bound energy state $|E_B\rangle$ can be generated from an unperturbed state $|E_0\rangle$ corresponding to energy E_0 by the formula

$$c|E_B\rangle = \frac{U_I(t_0, -\infty)|E_0\rangle}{\langle E_0|U_I(t_0, -\infty)|E_0\rangle} \qquad (10.232)$$

It happens, however, that the situation is slightly more complicated than this. For, in general, the unperturbed system will possess no stationary states of energy exactly E_0. Rather, the continuum levels will consist of all energy values greater than but not including E_0.

Only at the point $\lambda = \lambda^{\circ}$ will the system possess an energy level equal to E_0. While λ is still equal to λ_0 this level may be regarded as belonging to the continuum levels, but as soon as λ passes beyond this level slips over into a discrete level corresponding to a bound state. For example, in the discussion of the one-dimensional delta-function potential we might have chosen for our unperturbed Hamiltonian operator H_0 that corresponding to a system with a repulsive delta-function potential, the perturbation function H_1 being taken to be an over-compensating attractive delta-function potential. There would then have existed no unperturbed state with k = 0 (i.e. zero energy). To get around this difficulty the formula (10.232) for the bound state must be replaced by

$$c\,|E_B\rangle = \lim_{\epsilon \to +0} \frac{U_I(t_0, -\infty)\,|E_0 + \epsilon\rangle}{\langle E_0 + \epsilon | U_I(t_0, -\infty) | E_0 + \epsilon \rangle} \qquad (10.233)$$

The ϵ occurring in (10.233) is same ϵ which occurs in the adiabatic definition of $U_I(t_0, -\infty)$. Formula (10.233) holds even when $|E_0\rangle$ exists, and we could have carried out our discussion of the one-dimension. delta function potential with this improved formula if we had wished. (10.233) now gives us a slightly altered picture of the adiabatic formation of a bound state. The bound state is seen to come not from a singl possibly non-existent state $|E_0\rangle$, but rather from an infinitesimally thin layer of states lying just above E_0 in energy.

Suppose the parameter λ is gradually increased in magnitude beyond λ_0. In the general case a point, λ_1, will eventually be reached at which another bound state ill suddenly appear. And as λ is increased even further, more and more bound states may appear. (at points $\lambda_2, \lambda_3, \ldots$, etc.). Even an infinity of bound states may be possible* (e.g particle in a square well potential).

---

* In the case of Kepler motion (e.g. hydrogen atom) there are an infinity of bound states regardless of how small λ is taken. In order to treat such cases one must vary the perturbation function in some other manner than by nearly adjusting a multiplicative factor. For example, one might add a variable exponential "screening" factor).

In the example considered above, of the delta function potential, it just happened that only one bound state could be produced no matter how large λ was taken. In the general case, however, there may be more than one bound state. The adiabatic analysis of the situation involving several bound states can be carried out exactly as in the case of a single bound state. One must simply take care that only one bound state is allowed to come down at a time. For example, if we have the case of two bound states, with a λ lying between λ_1, and λ_2, we may denote this value of λ by λ_1', and choose another value of λ, say λ_0', lying between λ_α and λ_1. We must then first carry out the perturbation from $\lambda = 0$ to $\lambda = \lambda_0'$. This brings down one bound state. In carrying out the next step, from $\lambda = \lambda_0'$ to $\lambda = \lambda_1'$, the perturbed system at $\lambda = \lambda_0'$ is then to be regarded as the new unperturbed system. The passage to $\lambda = \lambda_1$ is effected by applying a perturbation proportional to $\lambda_1' - \lambda_0'$ to this sytem. This brings down the second bound state, while the bound state that is already present merely suffers a level shift, which can be treated by previously discussed methods. The first bound state will come from the states of the original unperturbed system which lie just above E_0 in energy. The second bound state will come from the states of the system at $\lambda = \lambda_0'$ which lie just above E_0 in energy./ However, in order to keep the two bound states separate one must take care that they are generated separately in the above fashion

It is fairly easy to show that the process of applying one perturbation after another is actually equivalent to applying the entire perturbation at once. Let the Hamiltonian function of the completely perturbed system be

$$H = H_0 + H_1 + H_2 \,. \tag{10.234}$$

Here we imagine that the perturbation H_1 is followed by the perturbation H_2. H_1 and H_2 may be completely arbitrary functions, not necessarily proportional to one another. Hence the derivation which follows is applicable to a wider class of problems than those for

/Hence this state too has its eventual origin in the states of the original unperturbed system which lie just above E_0 in energy.

which the perturbations consist simply in adjusting a single parameter λ.

The first perturbation is described by means of the operator (see (10.25) and (10.45))

$$U_I^{(1)}(t_0,-\infty)=P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{1}{i\hbar}H_{0S}(t_0-t)}H_{1S}e^{\frac{1}{i\hbar}H_{0S}(t-t_0)}dt\right], \quad (10.235)$$

Here the unperturbed Hamiltonian function is H_0. When the second perturbation is applied, however, the unperturbed Hamiltonian function must be taken to be $H_0 + H_1$. Hence the second perturbation is described by means of the operator

$$U_I^{(2)}(t_0,-\infty)=P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{1}{i\hbar}(H_{0S}+H_{1S})(t_0-t)}H_{2S}e^{\frac{1}{i\hbar}(H_{0S}+H_{1S})(t-t_0)}dt\right], \quad (10.236)$$

On the other hand, the overall perturbation is described by means of the operator

$$U_I(t_0,-\infty)=P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{1}{i\hbar}H_{0S}(t_0-t)}(H_{1S}+H_{2S})e^{\frac{1}{i\hbar}H_{0S}(t-t_0)}dt\right]. \quad (10.237)$$

In order to show that $U_I(t_0,-\infty)=U_I^{(2)}(t_0,-\infty)U_I^{(1)}(t_0,-\infty)$ we must first derive a theorem involving exponentials and the "P-bracket" or time-ordering operator.

Let $F(t)$ and $G(t)$ be any two time dependent operators, and consider the expression

$$P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\left[F(t)+G(t)\right]dt\right]. \quad (10.238)$$

Let us try to rewrite this expression in the form

$$P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\bar{F}(t)\,dt\right]P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right] \quad (10.239)$$

where the operator $\bar{F}(t)$ is independent of t'. Differentiating expression (10.238) with respect to t' and setting the result equal

to the same derivative of expression (10.239), we obtain

$$-P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}[F(t)+G(t)]\,dt\right][F(t')+G(t')]$$
$$=-P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\bar{F}(t)\,dt\right]\bar{F}(t')\,P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right]$$
$$-P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\bar{F}(t)\,dt\right]P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right]G(t') \quad (10.240)$$

Using the fact that (10.238) and (10.239) are to be regarded as equal, we obtain

$$-P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\bar{F}(t)\,dt\right]P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right]F(t')$$
$$=-P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}\bar{F}(t)\,dt\right]\bar{F}(t')\,P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right], \quad (10.241)$$

Remembering that the inverse of an operator of the form (10.238) is obtained simply by interchanging t' and t", we have, on multiplying eq. (10.241) on the left by $P\left[\exp\frac{1}{i\hbar}\int_{t''}^{t'}\bar{F}(t)\,dt\right]$ and on the right by $P\left[\exp\frac{1}{i\hbar}\int_{t''}^{t'}G(t)\,dt\right]$, and setting t' = t,

$$\bar{F}(t)=P\left[\exp\frac{1}{i\hbar}\int_{t}^{t''}G(t)\,dt\right]F(t)\,P\left[\exp\frac{1}{i\hbar}\int_{t''}^{t}G(t)\,dt\right] \quad (10.242)$$

Hence

$$P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}[F(t)+G(t)]\,dt\right]$$
$$=P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}P\left[\exp\frac{1}{i\hbar}\int_{t}^{t''}G(t''')\,dt'''\right]F(t)\,P\left[\exp\frac{1}{i\hbar}\int_{t''}^{t}G(t''')\,dt'''\right]dt\right]$$
$$\times P\left[\exp\frac{1}{i\hbar}\int_{t'}^{t''}G(t)\,dt\right] \quad (10.243)$$

Applying this theorem now to (10.237), we get

$$U_I(t_0,-\infty) = P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} P\left[\exp\frac{1}{i\hbar}\int_{t}^{t_0} e^{\frac{1}{i\hbar}H_{0S}(t_0-t''')} H_{1S} e^{\frac{1}{i\hbar}H_{0S}(t'''-t_0)}dt'''\right]\right.$$
$$\times e^{\frac{1}{i\hbar}H_{0S}(t_0-t)} H_{2S} e^{\frac{1}{i\hbar}H_{0S}(t-t_0)}$$
$$\left.\times P\left[\exp\frac{1}{i\hbar}\int_{t_0}^{t} e^{\frac{1}{i\hbar}H_{0S}(t_0-t''')} H_{1S} e^{\frac{1}{i\hbar}H_{0S}(t'''-t_0)}dt'''\right]dt\right]$$
$$\times P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{1}{i\hbar}H_{0S}(t_0-t)} H_{1S} e^{\frac{1}{i\hbar}H_{0S}(t-t_0)}dt\right] \quad (10.244)$$

Observing, however, that

$$e^{\frac{1}{i\hbar}(H_{0S}+H_{1S})(t_0-t)} = P\left[\exp\frac{1}{i\hbar}\int_{t}^{t_0}(H_{0S}+H_{1S})\,dt'''\right]$$
$$= P\left[\exp\frac{1}{i\hbar}\int_{t}^{t_0} P\left[\exp\frac{1}{i\hbar}\int_{t'''}^{t_0} H_{0S}\,dt'\right] H_{1S}\, P\left[\exp\frac{1}{i\hbar}\int_{t_0}^{t'''} H_{0S}\,dt'\right]dt'''\right]$$
$$\times P\left[\exp\frac{1}{i\hbar}\int_{t}^{t_0} H_{0S}\,dt'''\right]$$
$$= P\left[\exp\frac{1}{i\hbar}\int_{t}^{t_0} e^{\frac{1}{i\hbar}H_{0S}(t_0-t''')} H_{1S} e^{\frac{1}{i\hbar}H_{0S}(t'''-t_0)}dt'''\right] e^{\frac{1}{i\hbar}H_{0S}(t_0-t)}, \quad (10.245)$$

and recognizing that the last factor of (10.244) is simply $U_I^{(1)}(t_0,-\infty)$, we have

$$U_I(t_0,-\infty) = P\left[\exp\frac{1}{i\hbar}\int_{-\infty}^{t_0} e^{\frac{1}{i\hbar}(H_{0S}+H_{1S})(t_0-t)} H_{2S} e^{\frac{1}{i\hbar}(H_{0S}+H_{1S})(t-t_0)}dt\right] U_I^{(1)}(t_0,-\infty)$$
$$= U_I^{(2)}(t_0,-\infty)\,U_I^{(1)}(t_0,-\infty), \quad (10.246)$$

q.e.d.

The above results may obviously be generalized to the case of more than two bound states. As in the example of the delta-function potential considered above, the unitarity of the overall

transformation operator $U_I(t_0, -\infty)$ may be forcefully maintained by a purely formal procedure. Now, however, since all of the bound states have a common origin in the unperturbed energy levels just above E_0, they must be carefully traced back individually, each to its own infinitesimal unperturbed energy band. One must make a formal distinction between these energy bands, labelling them by some such scheme as $E_0^{(1)}$, $E_0^{(2)}$, $E_0^{(3)}$, ... etc., so that the bound states are given formally by

$$|E_B^{(i)}\rangle = U_I(t_0, -\infty)\,|E_0^{(i)}\rangle, \qquad i = 1, 2, 3 \ldots\ldots . \tag{10.247}$$

We pointed out, in the discussion of the one-dimensional delta-function potential, that the S-matrix was in that case unambiguously unitary in spite of the presence of a bound state. This result holds in general for all systems no matter how many bound states are present. We shall now proceed to verify this. Introducing the operators T_- and X_- defined by equations (10.136, 137, 138), we may write

$$\begin{aligned} 0 &= -2\pi i\, U_I^*(t_0, -\infty)\, H_{IS}\, U_I(t_0, -\infty) \\ &\quad + 2\pi i\, U_I^*(t_0, -\infty)\, H_{IS}\, U_I(t_0, -\infty) \\ &= (1 + T_-^*)\, X_- + X_-^*\,(1 + T_-) \\ &= X_- X_-^* + T_-^*\, X_- + X_-^*\, T_- \end{aligned} \tag{10.248}$$

Taking matrix elements of this operation, and remembering that $\langle \alpha_0'' | F^* | \alpha_0' \rangle = \langle \alpha_0' | F | \alpha_0'' \rangle^*$ and $\delta_-^* = \delta_+$, we obtain

$$\begin{aligned} 0 = {} & \langle \alpha_0'' | X_- + X_-^* | \alpha_0' \rangle \\ & + \sum_{\alpha_0'''} \left[\delta_+ (H_0''' - H_0'') + \delta_- (H_0''' - H_0') \right] \langle \alpha_0''' | X_-^* | \alpha_0''' \rangle \langle \alpha_0''' | X_- | \alpha_0' \rangle . \end{aligned} \tag{10.249}$$

Multiplying this equation by $\delta(H''_0 - H'_0)$, introducing the underlined operators defined by (10.140), and remembering that $\delta_+ + \delta_- = \delta$, we get

$$\underline{X}_- + \underline{X}^*_- + \underline{X}^*_- \underline{X}_- = 0 \qquad (10.250)$$

Remembering the form (10.242) for the S-operator, we see that (10.250) implies

$$S^* S = 1 \qquad (10.251)$$

This relation we might have expected as being implied already by (10.132). In order to prove the other necessary condition for unitarity, namely $S S^* = 1$, we must introduce the operator $U_I(\infty, t)$ which satisfies the differential equation

$$-i\hbar \frac{\partial}{\partial t} U_I(\infty, t) = U_I(\infty, t) H_{,I}(t) \qquad (10.252)$$

and the boundary condition

$$\lim_{t \to \infty} U_I(\infty, t) = 1 . \qquad (10.253)$$

Eqs. (10.252, 253) may be combined into the single integral equations

$$U_I(\infty, t) = 1 + \frac{1}{i\hbar} \int_t^\infty U_I(\infty, t') H_{,I}(t')\, dt' \qquad (10.254)$$

In order to express this operator in a different form let us set

$$U_I(\infty, t) = e^{-\frac{1}{i\hbar} H_0 (t - t_0)} U_I(\infty, t_0) e^{\frac{1}{i\hbar} H_0 (t - t_0)} \qquad (10.255)$$

in analogy with (10.56). Then, using (10.45), we may write (cf. (10.61))

$$U_I(\infty,t_0) = 1 + \frac{1}{i\hbar}\int_{t_0}^{\infty} e^{-\frac{1}{i\hbar}H_0(t'-t_0)}\, U_I(\infty,t_0)\, H_{1S}\, e^{\frac{1}{i\hbar}H_0(t'-t_0)}\, dt' \qquad (10.256)$$

$$= 1 + T_+$$

where

$$\langle \alpha''_0 | T_+ | \alpha'_0 \rangle = \delta_+(H''_0 - H'_0)\langle \alpha''_0 | X_+ | \alpha'_0 \rangle \qquad (10.257)$$

with

$$X_+ = -2\pi i\, U_I(\infty, t_0)\, H_{1S} \qquad (10.258)$$

Proceeding exactly as we did for the operators X_- and T_-, we may write

$$\begin{aligned} 0 &= 2\pi i\, U_I(\infty,t_0)\, H_{1S}\, U_I^*(\infty,t_0) \\ &\quad - 2\pi i\, U_I(\infty,t_0)\, H_{1S}\, U_I^*(\infty,t_0) \\ &= (1+T_+)X_+^* + X_+(1+T_+^*) \\ &= X_+ + X_+^* + X_+ T_+^* + T_+ X_+^* \end{aligned} \qquad (10.259)$$

Taking matrix elements of this equation, and multiplying by $\delta(H''_0 - H'_0)$, we get

$$\underline{X}_+ + \underline{X}_+^* + \underline{X}_+\underline{X}_+^* = 0 . \qquad (10.260)$$

But we also have

$$\begin{aligned} 0 = &- 2\pi i\, U_I(\infty,t_0)\, H_{1S}\, U_I(t_0,-\infty) \\ &+ 2\pi i\, U_I(\infty,t_0)\, H_{1S}\, U_I(t_0,-\infty) \end{aligned}$$

$$= \Lambda_+ (1 + T_-) - (1 + T_+) \Lambda_-$$

$$= \Lambda_+ - \Lambda_- + \Lambda_+ T_- - T_+ \Lambda_- \,. \qquad (10.261)$$

Taking matrix elements of this equation, we get

$$0 = \langle \alpha''_0 | \Lambda_+ - \Lambda_- | \alpha'_0 \rangle$$

$$+ \sum_{\alpha \neq 0} \left[\delta_- (H'''_\alpha - H'_0) - \delta_+ (H''_0 - H'''_\alpha) \right] \langle \alpha'_\alpha | \Lambda_+ | \alpha''_0 \rangle \langle \alpha'''_0 | \Lambda_- | \alpha'_\alpha \rangle \qquad (10.262)$$

Multiplying (10.262) by $\delta(H''_0 - H'_0)$ and remembering that $\delta_-(-x) = \delta_+(x)$ we get

$$\underline{\Lambda}_+ = \underline{\Lambda}_- \qquad (10.263)$$

Hence, from (10.260)

$$\underline{\Lambda}_- + \underline{\Lambda}^*_- + \underline{\Lambda}_- \underline{\Lambda}^*_- = 0, \qquad (10.264)$$

which implies

$$S S^* = 1 \qquad (10.265)$$

The fact that $\underline{\Lambda}_+ = \underline{\Lambda}_-$ may also be inferred from another argument. The S-operator should be definable by the equation

$$S = U_I(\infty, -\infty) = 1 + \frac{1}{i\hbar} \int_{-\infty}^{\infty} U_I(\infty, t) H_{int}(t) \, dt \qquad (10.266)$$

equally well as by the equation (10.142). But substituting (10.45) and (10.255) into (10.266), we get (10.142))

$$S = 1 + \frac{1}{i\hbar}\int_{-\infty}^{+\infty} e^{-\frac{1}{i\hbar}H_0(t'-t_0)}\, U_I(\infty,t_0)\, H_{1S}\, e^{\frac{1}{i\hbar}H(t'-t_0)}\, dt'$$

$$= 1 + \underline{X}_+ \qquad (10.267)$$

from which (10.263) follows.

It will be instructive to present here another proof of the unitarity of the S-matrix by a method which is very useful in the theory of scattering. Let us go all the way back to the integral equation (10.63) and rewrite it in the form

$$|\alpha'_0\rangle = |\alpha'_0\rangle - \left[\frac{1}{H_0 - H'_0} + \pi i\,\delta(H_0 - H'_0)\right] H_{1S}\,|\alpha'_0\rangle \qquad (10.268)$$

in which it is understood that the principal value is to be taken of any integral in which the singular function $\frac{1}{H_0 - H'_0}$ appears. If we introduce the vector

$$|\alpha'_0\rangle^{\text{shuffle}} = |\alpha'_0\rangle - \pi i\,\delta(H_0 - H'_0)\, H_{1S}\,|\alpha'_0\rangle$$

$$= \left(1 + \frac{1}{2}\underline{X}\right)|\alpha'_0\rangle \qquad (10.269)$$

where

$$\underline{X} \equiv \underline{X}_+ \equiv \underline{X}_- \qquad (10.270)$$

we may write

$$|\alpha'_0\rangle = \left[1 + \frac{1}{H_0 - H'_0} H_{1S}\right]^{-1} |\alpha'_0\rangle^{\text{shuffle}} \qquad (10.271)$$

Equations (10.269) and (10.271) have an interesting physical significance. First observe that if there were no degeneracy in the unperturbed system there would be no scattering. For, remember that the S-operator, like the operators $U_I(t_0,-\infty)$ and $U_I(\infty,t_0)$, must reduce

be defined by means of an adiabatic prescription. The delta function of the energy differences, which occurs in the S-matrix, arises only because of the adoption of some procedure such as introducing exponential "dumping" factors (cf. the analysis of the delta function in the Appendix). If we write the S-operator in the form

$$S = U_I(\infty, -\infty) = U_I(\infty, t_0)\, U_I(t_0, -\infty) \qquad (10.272)$$

then we have an explicit characterization of it as the operator which switches the perturbation on, and then off again, adiabatically. If none of the unperturbed energy states is degenerate then we know, by the theorem of adiabatic invariance, that the system must eventually find itself in the same state as that in which it started. If, on the other hand, the unperturbed states are degenerate, the switching on and off of the perturbation produces, in general, a shuffling of the degenarate states corresponding to a given energy level, and the net result appears physically as a scattering. The final state is no longer required to be the same as the initial state. It is only required to have the same energy [*].

Equation (10.269) is an explicit expression of the shuffling process. The vector $|\alpha'_0\rangle_{shuffle}$ expresses the extent to which the initial state $|\alpha'_0\rangle$ is shuffled in transforming to the perturbed stationary state $|\alpha'\rangle$. It should be observed that since $H_0\,\delta(H_0 - H'_0) = H'_0\,\delta(H_0 - H'_0)$ the vector $|\alpha'_0\rangle_{shuffle}$ is also an eigenvector of the operator H_0 corresponding

---

[*] For certain systems other conservation laws besides that of energy may be important. In field theories, for example, the final state must have also the same momentum, angular momentum, and charge as the initial state.

ponding to the eigenvalue H'_0 :

$$H_0 \left| \right\rangle^{shuffle}_{\alpha'_0} = H'_0 \left| \right\rangle^{shuffle}_{\alpha'_0} \tag{10.273}$$

The vector $\left| \right\rangle_{\alpha'_0}$ can be obtained from the vector $\left| \right\rangle^{shuffle}_{\alpha'_0}$ by applying a pure "strain", as represented by the operator $\left[1 + \frac{1}{H_0 - H'_0} H_{1S}\right]^{-1}$ in (10.271). It is this operator which puts the detailed structure into the perturbed stationary state. The above procedure may also be reversed. That is, we may apply the "strain" to the initial states first, and then let them be shuffled. Thus, if we write

$$\left| \right\rangle^{strain}_{\alpha'_0} = \left[1 + \frac{1}{H_0 - H'_0} H_{1S}\right]^{-1} \left| \alpha'_0 \right\rangle \tag{10.274}$$

we have, on substituting (10.269) into (10.271), and making use of (10.274)

$$\left| \right\rangle_{\alpha'_0} = \left| \right\rangle^{strain}_{\alpha'_0} + \frac{1}{2} \sum'_{\alpha''_0} \left| \right\rangle^{strain}_{\alpha''_0} \left\langle \alpha''_0 \right| \underline{X} \left| \alpha'_0 \right\rangle \tag{10.275}$$

Let us now introduce the operator

$$K = \sum_{\alpha'_0} H_{1S} \left[1 + \frac{1}{H_0 - H'_0} H_{1S}\right]^{-1} \left| \alpha'_0 \right\rangle \left\langle \alpha'_0 \right| \tag{10.276}$$

whose matrix elements may, in virtue of (10.274), be written in the form

$$\left\langle \alpha''_0 \right| K \left| \alpha'_0 \right\rangle = \left\langle \alpha''_0 \right| H_{1S} \left| \right\rangle^{strain}_{\alpha'_0} \tag{10.277}$$

The underlined form of the operator K, namely

$$\underline{K} = \sum_{\alpha'_0} \delta (H_0 - H'_0) H_{1S} \left[1 + \frac{1}{H_0 - H'_0} H_{1S}\right]^{-1} \left| \alpha'_0 \right\rangle \left\langle \alpha'_0 \right| \tag{10.278}$$

is sometimes known as the "reactance" operator.

Using (10.275) and (10.277) we may write

$$\langle \alpha''_0 | X | \alpha'_0 \rangle = -2\pi i \langle \alpha''_0 | H_{1S} U_I(t_0, -\infty) | \alpha'_0 \rangle$$

$$= -2\pi i \langle \alpha''_0 | H_{1S} | \rangle^{\text{strain}}_{\alpha'_0} - \pi i \sum_{\alpha'''_0} \langle \alpha''_0 | H_{1S} | \rangle^{\text{strain}}_{\alpha'''_0} \langle \alpha'''_0 | X |$$

$$= -2\pi i \langle \alpha''_0 | K | \alpha'_0 \rangle \qquad (10.279)$$

$$- \pi i \sum \langle \alpha''_0 | K | \alpha'''_0 \rangle \langle \alpha'''_0 | X | \alpha'_0 \rangle$$

Multiplying (10.279) by $\delta(H''_0 - H'_0)$ we get

$$\underline{X} = -2\pi i \underline{K} - \pi i K X \qquad (10.280)$$

Equation (10.280) is known as the Heitler integral equation. Its solution may be written down formally as

$$\underline{X} = \frac{-2\pi i \underline{K}}{1 + \pi i \underline{K}} \qquad (10.281)$$

Finally we may express the S-operator in terms of $\underline{K}$:

$$S = 1 + \underline{X} = \frac{1 - \pi i \underline{K}}{1 + \pi i \underline{K}} \qquad (10.282)$$

If we can show that the operator $\underline{K}$ is Hermitian then it follows that S is unitary. The demonstration that $\underline{K}$ is Hermitian is quite easy. First observe that $\underline{K}$ can be written in the forms

$$\underline{K} = \sum_{\alpha'_0} \delta(H_0 - H'_0) K(H'_0) | \alpha'_0 \rangle \langle \alpha'_0 |$$

$$= \sum_{\alpha''_0} | \alpha''_0 \rangle \langle \alpha''_0 | K(H''_0) \delta(H''_0 - H_0) \qquad (10.283)$$

where

$$K(H_0') = H_{1S}\left[1 + \frac{1}{H_0 + H_0'} H_{1S}\right]^{-1} \tag{10.284}$$

If we can show that the operators $K(H_0')$ are Hermitian then it follows /then that $\underline{K}$ is Hermitian, since/the two forms in (10.283) will be the Hermitian adjoints of one another. Now, we have

$$K^*(H_0') = \left[1 + H_{1S}\frac{1}{H_0 - H_0'}\right]^{-1} H_{1S} \tag{10.285}$$

and hence

$$K^*(H_0') = \left[1 + \frac{1}{H_0 - H_0'} H_{1S}\right]^{-1} = \left[1 + H_{1S}\frac{1}{H_0 - H_0'}\right]^{-1} H_{1S}\left[1 + \frac{1}{H_0 - H_0'} H_{1S}\right]$$

$$= \left[1 + H_{1S}\frac{1}{H_0 - H_0'}\right]^{-1}\left[1 + H_{1S}\frac{1}{H_0 - H_0'}\right] H_{1S} = H_{1S} \tag{10.286}$$

Multiplying eq. (10.286) on the right by $\left[1 + \frac{1}{H_0 - H_0'} H_{1S}\right]^{-1}$ we obtain

$$K^*(H_0') = H_{1S}\left[1 + \frac{1}{H_0 - H_0'} H_{1S}\right]^{-1} = K(H_0') \tag{10.287}$$

q.e.d.

APPENDIX . - CALCULATION OF CONSTANT a .

We can determine the value of "a" by a wave packet argument. To do this let us first introduce a notation which is reminiscent of the treatment of the vector potential A_μ in chapter 3. In chapter 3 we split the vector potential into "incoming" and "retarded", or "outgoing" and "advanced" parts. Let us now perform a similar separation of the state vector of our scattering problem. Let $|\ \rangle$ denote the Heisenberg state vector of the system. ($|\ \rangle$ does not necessarily have to be a stationary state. Introduce a coordinate representation for the system, and then examine the asymptotic behavior of the wave function $\langle q'|\ \rangle$ for large values of the coordinates q. Let us write

$$\langle q'|\ \rangle \sim \langle q'|\ \rangle^{in} + \langle q'|\ \rangle^{ret} \qquad (10.141a)$$

where the notation will become clearer presently. Since we shall nowhere make use of the special properties of the system of coordinates q, we may, without ambiguity, also simply write

$$|\ \rangle \sim |\ \rangle^{in} + |\ \rangle^{ret} \qquad (10.141b)$$

We shall now show that the "incoming" state-vector $|\ \rangle^{in}$ is the unperturbed state vector from which $|\ \rangle$ is generated through application of the operator $U_I(t_0, -\infty)$. We shall do this by first assuming it true and then demonstrating the self consistency of the assumption. In particular then, if $|\ \rangle$ is equal to $|\ \rangle_{\alpha_0'}$, $|\ \rangle^{in}$ is equal to $|\alpha_0'\rangle$. We write

$$|\ \rangle_{\alpha_0'} \sim |\alpha_0'\rangle + |\ \rangle^{ret}_{\alpha_0'} . \qquad (10.141c)$$

Construct next a wave packet out of the states $|\ \rangle_{\alpha_0'}$;

$$|\ \rangle = \sum_{\alpha'_0} f(\alpha'_0)|\ \rangle_{\alpha'_0} \qquad (10.141d)$$

where f is a function which is "peaked" around some particular set of values of the α'_0. Hence H' = H'o in the present problem (i.e. there is no level shift) the Schrödinger state vector corresponding to (10.141d) has the form

$$|t\rangle = e^{-i/\hbar H(t-t_0)}|\ \rangle$$

$$= \sum_{\alpha'_0} e^{-i/\hbar H'_0(t-t_0)} f(\alpha'_0)|\ \rangle_{\alpha'_0} \qquad (10.141e)$$

Strictly speaking, it is the "moving" function $\langle q'|t\rangle$ which describes the "moving" wave packet.

Let us now examine the properties of $|t\rangle$ as $t \to -\infty$. If the wave packet is arranged as to be an incoming packet, then, for large negative values of the time, the packet will still be at a large distance from the scattering center. At such times only the asymptotic parts of the vectors $|\ \rangle_{\alpha'_0}$ will contribute essentially to building up the wave packet, all other parts mutually cancelling owing to rapidly varying phases. Thus we may write

$$\lim_{t\to-\infty} |t\rangle = \lim_{t\to-\infty} \sum_{\alpha'_0} e^{-i/\hbar H'_0(t-t_0)} f(\alpha'_0) \langle \alpha'_0|q'\rangle \langle q'|\ \rangle^{ret}_{\alpha'_0} \qquad (10.141f)$$

Although the functions $\langle q'|\alpha'_0\rangle$ and $\langle q'|\ \rangle^{ret}_{\alpha'_0}$ are defined at points close in to the scattering center (as well as at distant points), these regions will not contribute to the wave packet, likewise on account of the mutual cancellation due to rapidly varying phases.

It should now be observed that although $|\alpha'_0\rangle$ is an eigenvector of Ho, the vector $|\ \rangle^{ret}_{\alpha'_0}$ is not. More precisely, the function

$$\langle p'|H_0 - H'_0|\ \rangle^{ret}_{\alpha'_0}$$

vanishes at all points q' except at the "origin" (i.e. at the scattering center). This condition plus the boundary conditions imposed by the asymptotic form of $\langle q' | \rangle^{ret}_{\alpha_0}$ are, in fact, sufficient to determine $| \rangle^{ret}_{\alpha_0}$ uniquely. The vector $| \rangle^{ret}_{\alpha_0}$ is therefore a kind of state-vector generated by a point source located at the origin. The situation here is analogous to that encountered in chapter 3., in which the vector A^{ret}_{μ} fails to satisfy the free wave function by a term involving the current source j_μ.

It should next be observed that

$$\lim_{t \to -\infty} \sum_{\alpha_0} e^{i/\hbar H_0 (t-t_0)} f(\alpha_0) | \rangle^{ret}_{\alpha_0} \qquad (10.141g)$$

because the function $\sum_{\alpha_0} e^{i/\hbar H_0'(t-t_0)} f(\alpha_0) \langle q' | \rangle^{ret}_{\alpha_0}$ represents an "outgoing" wave packet, and in the limit of large negative times, this outgoing packet has not yet emerged from the origin. Or, if one prefers to look at it in a different way, under time reversal the packet appears to be "sucked in" by the origin, and disappears. Owing to the point source generation of the vectors $| \rangle^{ret}_{\alpha_0}$, probability is not conserved for wave functions constructed out of these vectors alone.

Equation (10.141g) expresses the crucial point of the argument and shows why the use of wave packets is absolutely necessary. For, if the wave function $\sum_{\alpha_0} e^{i/\hbar H_0 (t-t_0)} f(\alpha_0) \langle q' | \rangle^{ret}_{\alpha_0}$ were spread out over all space, it would never get completely sucked into the origin under time reversal. Moreover, we could not write eq. (10.141f) in the first place, as non-asymptotic parts of the $| \rangle_{\alpha_0}$ would always contribute to $| \rangle$. These are actually very delicate points, and in a rigorous analytical discussion of our problem they must be handled very carefully. In reality, for example, no wave packet can be truly finite in a non-relativistic theory. It will always have at least a thin (but long) tail extending over all space. One must therefore make sure that this tail always has such negligible effect that none of the preceding arguments are basically invalidated.

(For a discussion of this difficulty, and the whole problem of non-relativistic scattering, see G.N. van Kampen, Phys. Rev.).

In virtue of eq. (10.141g) we may now write

$$\lim_{t\to-\infty} |t\rangle = \lim_{t\to-\infty} \sum_{\alpha_0'} e^{\frac{1}{i\hbar}H_0'(t-t_0)} f(\alpha_0')\,|\alpha_0'\rangle \qquad (10.141h)$$

That is, while the packet is still at a large distance from the scattering center it can be built out of the stationary state wave functions of the unperturbed system equally well as out of the corresponding stationary wave functions of the perturbed system. Moreover, while it is still at a large distance it doesn't matter very much whether the scattering perturbation H_1 is present or not. Hence H_1 can be switched on adiabatically while the wave packet approaches, and all of the perturbation-theoretic results obtained previously by adiabatic procedure are available to us.

When the packet gets near to the scattering center, then of course, it becomes necessary to use the perturbed wave functions. However, after the packet has been scattered (it generally gets broken up in the process and is moving away from the scattering center, then the use of the unperturbed wave functions can be resumed. The unperturbed wave functions used after the scattering process will not be those used prior to the scattering process, but will be connected with the latter by some transformation law.

To see this, let us introduce a set of vectors $|\,\rangle^{adv}_{\alpha_0}$ which, together with the vectors $|\,\rangle^{ret}_{\alpha_0}$, satisfy

$$(H_0 - H_0')\left(|\,\rangle^{ret}_{\alpha_0} - |\,\rangle^{adv}_{\alpha_0}\right) = 0 \qquad (10.141i)$$

and

$$\lim_{t\to\infty} \sum_{\alpha_0'} e^{\frac{1}{i\hbar}H_0'(t-t_0)} f(\alpha_0')\,|\,\rangle^{adv}_{\alpha_0'} = 0 \qquad (10.141j)$$

In practice, the vectors $|\,\rangle^{adv}_{\alpha_0}$, which are uniquely defined by

(10.141i, j) can be determined almost by inspection from the vectors $|\rangle^{ret}_{\alpha_0}$. The function $\sum_{\alpha_0} e^{i/\hbar H_0'(t-t_0)} f(\alpha_0) \langle \varphi' | \rangle^{ret}_{\alpha_0}$ corresponds to an "incoming" wave, and, in fact, is just such as to compensate for the lack of probability conservation in the "outgoing" wave $\sum_{\alpha_0} e^{i/\hbar H_0'(t-t_0)} f(\alpha_0) \langle \varphi' | \rangle^{ret}_{\alpha_0}$.
If we define

$$|\rangle^{rad}_{\alpha_0} = |\rangle^{ret}_{\alpha_0} - |\rangle^{adv}_{\alpha_0} \qquad (10.141k)$$

we may write

$$H_0 |\rangle^{rad}_{\alpha_0} = H_0' |\rangle^{rad}_{\alpha_0} \qquad (10.141L)$$

and we see that $\langle \varphi' | \rangle^{rad}_{\alpha_0}$ is a stationary state wave function of the unperturbed system.

Now, referring back to eqs. (10.141b, c, d) it is evident that we may write

$$|\rangle^{in} = \sum_{\alpha_0} f(\alpha_0) |\alpha_0\rangle \qquad (10.141m)$$

$$|\rangle^{ret} = \sum_{\alpha_0} f(\alpha_0) |\rangle^{ret}_{\alpha_0} \qquad (10.141n)$$

Let us also define

$$|\rangle^{adv} = \sum_{\alpha_0} f(\alpha_0) |\rangle^{adv}_{\alpha_0} \qquad (10.141o)$$

If, following the analogy of chapter 8, we express the asymptotic behavior of the vector $|\rangle$ also in the form*

$$|\rangle \sim |\rangle^{out} + |\rangle^{adv} \qquad (10.141p)$$

we may write

$$|\rangle^{out} = |\rangle^{in} + |\rangle^{rad} \qquad (10.141q)$$

where

$$|\,\rangle^{rad} = \sum_{\alpha_0} f(\alpha_0')\,|\,\rangle_{\alpha_0}^{rad} \qquad (10.141r)$$

The vectors $|\,\rangle^{out}$ and $|\,\rangle^{in}$ are connected by an operator S which, in its matrix form, is known as the S-matrix.

$$|\,\rangle^{out} = S\,|\,\rangle^{in} \qquad (10.141s)$$

In the limit of a very broad wave packet (although, of course, we must never allow the packet to become infinitely broad) the function $f(\alpha_0')$ becomes very sharply peaked, and equation (10.141s) effectively reduces to

$$S|\alpha_0'\rangle = |\alpha_0'\rangle + |\,\rangle_{\alpha_0'}^{rad} \qquad (10.141t)$$

(10.141t) may be taken as the defining equation of the S-matrix. The matrix elements are evidently given by

$$\langle \alpha_0''|S|\alpha_0'\rangle = \delta_{\alpha_0''\alpha_0'} + \langle \alpha_0''|\,\rangle_{\alpha_0'}^{rad} \qquad (10.141u)$$

☆ We could have carried out all of this analysis replacing (10.141d) and (10.141p) by true rather than asymptotic equations : i.e.

$|\,\rangle = |\,\rangle^{in} + |\,\rangle^{ret} = |\,\rangle^{out} + |\,\rangle^{radiation}$ ☆☆

This would have been closer to the spirit of eqs. (3.) and (3.) of chapter 3. However, we have chosen the present method to emphasize that the S-matrix (and hence scattering) can be analyzed completely in terms of the asymptotic behavior of the stationary state wave functions. The discussion of chapter 3, on the other hand, could not be carried out (except in first approximation) in terms of the asymptotic behavior of the vector potential A_μ, since this asymptotic behavior is not constant in time owing to the fact that the scattering particle moves and is gradually accelerated by the incoming field.

☆☆ The equations $\langle p'|H_0 - H_0'|\,\rangle_{\alpha_0}^{ret} = 0$, $\langle p'|H_0 - H_0'|\,\rangle_{\alpha_0}^{radiation} =$ would then have failed over a diffuse region (the region occupied by the scattering potential) rather than at a single point.

and may be read off directly by inspection of the asymptotic form of the stationary state perturbed wave function $\langle \varphi' |\rangle_{\alpha_0}$.

Let us now see how to describe scattering in terms of the S-matrix. Since the distant scattered wave packet may be described by the asymptotic wave functions, we may write

$$\lim_{t\to\infty} |t\rangle = \lim_{t\to+\infty} \sum_{\alpha_0'} e^{-\frac{i}{\hbar}H_0(t-t_0)} f(\alpha_0') \left(|\alpha_0'\rangle + |\rangle_{\alpha_0}^{ret} \right) \tag{10.141v}$$

In virtue of (10.141j) we may also write

$$\lim_{t\to\infty} |t\rangle = \lim_{t\to\infty} \sum_{\alpha_0'} e^{-\frac{i}{\hbar}H_0(t-t_0)} f(\alpha_0') \left(|\alpha_0'\rangle + |\rangle_{\alpha_0}^{rad} \right) \tag{10.141w}$$

Equation (10.141w) is an explixit expression of the scattered packet in terms of unperturbed stationary state wave functions.

It will now be convenient to introduce the interaction representation. Making use of equations (10.17), (10.26) and (10.141h), we have

$$|-\infty\rangle_I = \lim_{t\to-\infty} e^{\frac{i}{\hbar}H_0(t_0-t)} |t\rangle = \sum_{\alpha_0'} f(\alpha_0') |\alpha_0'\rangle \tag{10.141x}$$

Making use of (10.141w), we also have

$$|+\infty\rangle_I = \lim_{t\to\infty} e^{\frac{i}{\hbar}H_0(t_0-t)} |t\rangle = \sum_{\alpha_0'} f(\alpha_0') \left(|\alpha_0'\rangle + |\rangle_{\alpha_0}^{rad} \right) \tag{10.141y}$$

But, on account of (10.141t), (10.141y) may be written

$$|\infty\rangle_I = \sum_{\alpha_0'} f(\alpha_0') S |\alpha_0'\rangle = S |-\infty\rangle_I \tag{10.141z}$$

But

$$|\infty\rangle_I = U_I(\infty, -\infty) |-\infty\rangle_I \tag{10.141aa}$$

Hence we may make the identification*

$$S \triangleq U_I(\infty, -\infty) \tag{10.141bb}$$

and it is evident that in order to calculate the operator S we may avail ourselves of the computational techniques which have been developed previously for evaluating the operators $U_I(t'', t')$, in addition to the technique of examining asymptotic wave functions.

The process of scattering is completely described by the operator S. It is for this reason that S is often referred to as the scattering operator. From (10.45), (10.56) and (10.60) we find that the scattering operator is given by

$$\begin{aligned} S &= 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} H_{IS}(t')\, U_I(t', -\infty)\, dt' \\ &= 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} e^{-\frac{1}{i\hbar}H_0(t'-t_0)} H_{IS}\, U_I(t', -\infty)\, e^{\frac{1}{i\hbar}H_0(t'} \\ &= 1 + \frac{1}{i\hbar}\sum_{\alpha_0}\int_{-\infty}^{\infty} e^{-\frac{1}{i\hbar}(H_0 - H_0')t'} dt\, H_{IS}\, U_I(t', -\infty)\, |\alpha_0'\rangle\langle \\ &= 1 - 2\pi i \sum_{\alpha_0'} \delta(H_0 - H_0')\, H_{IS}\, U_I(t', -\infty)\, |\alpha_0'\rangle\langle\alpha_0'| \\ &= 1 + \underline{\chi}_- . \end{aligned} \tag{10.142}$$

The constant a mentionned above on pages X-363 is thus seen to be equal to unity.

---

* Evidently the vectors $|\ \rangle^{in}$ and $|\ \rangle^{out}$ may be identified with $|-\infty\rangle_I$ and $|\infty\rangle_I$ respectively, even to the phase factors.

QUANTUM MECHANICS

COURS DE L'ECOLE D'ETE DE PHYSIQUE THEORIQUE

QUANTUM MECHANICS

Bryce S. DE WITT

APPENDIX

LES HOUCHES
ETE 1953

A P P E N D I X

1. The delta function.

The "delta function" is a symbol $\delta(x)$ which satisfies the following formal equation :

$$\int_{-a}^{+a} f(x)\,\delta(x)\,dx = f(0) \qquad \text{for all } a>0 \text{ and all } f(x) \tag{1.1.}$$

where $f(x)$ is any arbitrary function which is continuous at the origin. The delta function defines a "mapping" of the set of all continuous functions into the straight line. Each continuous function is mapped into its value at the origin. Following L. Schwartz, the notion of "distribution" (a generalization of the notion of "function") may be introduced in terms of such mappings, and the formal properties of the delta function may be mathematically rigorized.

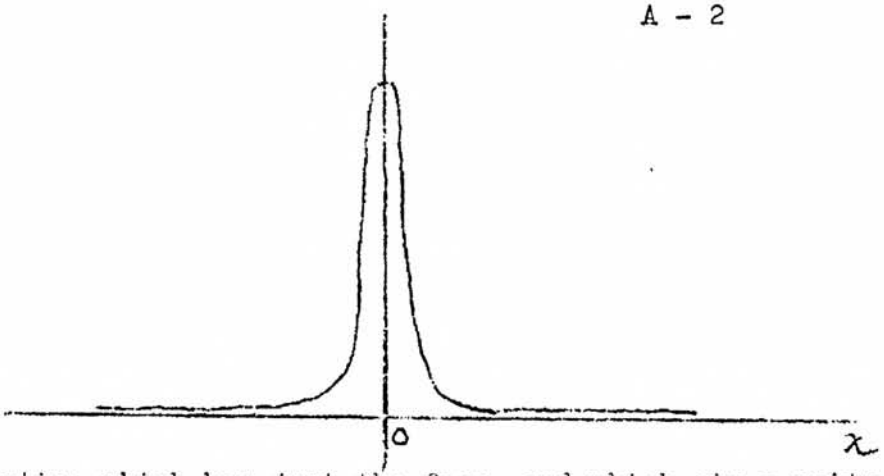
For such rigor the reader is referred to the literature. In what follows we shall stick close to the purely formal development of the theory of the delta function. This procedure has the advantage of being graphic, of often affording physical or pictorial insight, and of enabling us to stay close to the most widely used notations of contemporary theoretical physicists.

It will be observed from (1.1) that, since a and $f(x)$ are completely arbitrary, we may write, in a formal sense

$$\delta(x) = 0 \qquad \text{for } x \neq 0 \tag{1.2}$$

$$\int_{-\infty}^{+\infty} \delta(x) = 1 \tag{1.3}$$

These relations, although purely formal, give us a mental picture of the delta function as a sort of limiting form of a very steep function which has the following appearance.



A function which has just the form, and which gives unity when integrated from $-\infty$ to $+\infty$, is

$$\frac{1}{\pi}\ \frac{\epsilon}{x^2+\epsilon^2}\ , \tag{1.4}$$

where ϵ is a small positive number.

The delta-function nature of this function is readily seen through consideration of the following integral

$$\frac{1}{\pi}\int_{-a}^{+a} f(x)\ \frac{\epsilon}{x^2+\epsilon^2}\ dx = \frac{1}{\pi}\left[f(x)\ \tan^{-1}\frac{x}{\epsilon}\right]_{-a}^{+a} - \frac{1}{\pi}\int_{-a}^{+a} f'(x)\ \tan^{-1}\frac{x}{\epsilon}\ dx$$

Passing to the limit as $\epsilon \to 0$, and using the fact that

$$\lim_{\epsilon\to 0}\ \frac{1}{\pi}\ \tan^{-1}\frac{x}{\epsilon} = \frac{1}{2}\ \frac{x}{|x|}\ , \tag{1.5}$$

we have

$$\lim_{\epsilon\to 0}\int_{-a}^{+a} f(x)\frac{\epsilon}{x^2+\epsilon^2}\ dx = \frac{1}{2}\left[f(a)\frac{x}{|x|}\right]_{-a}^{+a} - \frac{1}{2}\int_{-a}^{+a} f'(x)\ \frac{x}{|x|}\ dx$$

$$= \frac{1}{2}\left[f(a)+f(-a)\right] + \frac{1}{2}\int_{-a}^{0} f'(x)\,dx - \frac{1}{2}\int_{0}^{a} f'(x)\,dx$$

$$= \frac{1}{2}\left[f(a)+f(-a)\right] + \frac{1}{2}\left[f(0)-f(-a)\right] - \frac{1}{2}\left[f(a)-f(0)\right]$$

$$= f(0) \tag{1.6}$$

Thus the delta function can be replaced by the function (1.4) in any integral expression, provided the passage to the limit $\epsilon \to 0$ is carried out after the integration has been performed. In fact, such a procedure of passage to a limit is actually implied by the formal expression (1.1)

Henceforth, whenever we deal with expressions involving an ϵ, it will always be understood that a limiting process $\epsilon \to 0$ is to be carried out after all the analytical operations involving such expressions have been performed. With this "implicit limiting convention", then, we may write

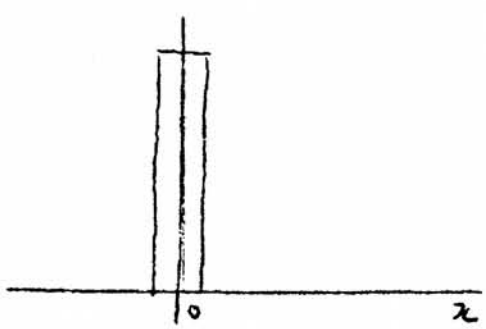
$$\delta(x) = \frac{1}{\pi} \frac{\epsilon}{x^2+\epsilon^2} \tag{1.7}$$

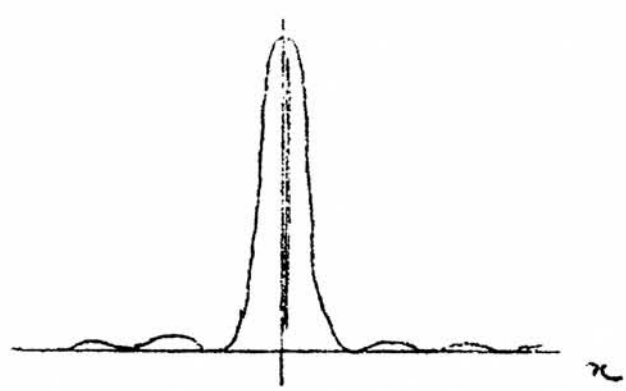
From (1.7) we obtain

$$\delta(0) = \frac{1}{\pi\epsilon} \to \infty \tag{1.8}$$

That is, $\delta(0)$ is a very large number, which becomes infinite in the limit. In spite of the divergent nature of $\delta(x)$ at the origin, we shall on occaision find it convenient to make use of the expression $\delta(0)$ in our anylitical work. Sometimes, in order to make the work more palatable for some (or more graphic for others), the symbol $\delta(0)$ will be replaced by an equivalent symbol such as $1/\pi\epsilon$.

Expression (1.4) is not the only function which possesses a delta-function behavior in the limit. For example, functions which have general behaviors of the following types could equally well be used to define the delta function by a limiting process :





However, the form (1.4) happens to be particularly useful and will suffice for a very large part of our analytical work with the delta function. Indeed, it will appear that with the use of (1.4) and equivalent expressions, a large part of that portion of analysis which is related to Fourier transform theory can be carried out in terms of the elementary functions alone.

From (1.7) it appears that the delta function is an even function. This can also be verified directly in the following manner.

$$\int_{-a}^{+a} f(x)\,\delta(-x)\,dx = -\int_{a}^{-a} f(-y)\,\delta(y)\,dy \qquad (y=-x)$$

$$= \int_{-a}^{+a} f(-y)\,\delta(y)\,dy = f(0) \qquad (1.9)$$

$\delta(-x)$ is seen to have precisely the same effect as $\delta(x)$, and is therefore, to all intents and purposes, indistinguishable from $\delta(x)$

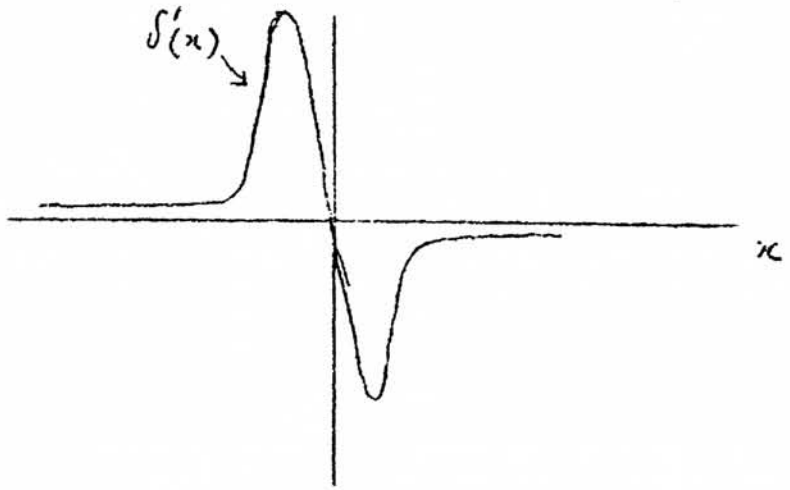
$$\delta(-x) = \delta(x) \qquad (1.10)$$

The "derivative of the delta function", on the other hand, is an odd function of :

$$\delta'(x) = -\frac{2}{\pi}\,\frac{\epsilon x}{(x^2+\epsilon^2)^2} \qquad (1.11)$$

Before passage to the limit it has the following form : (see graph on next page.)

Since $\delta(x) = 0$ for $x \neq 0$, it is evident that $\delta'(x) = 0$ for points x not in the neighborhood of the origin. It should be noted however, that it is not quite proper to say " $\delta'(x) = 0$ for $x \neq 0$". For this combined with $\delta'(0) = 0$, which follows from the odd nature of $\delta'(x)$, would imply $\delta'(x) = 0$ everywhere. Instead we should investigate the limiting process a little more carefully, and this applies also to the delta function itself.



Suppose, using expression (1.7), we ask ourselves how far from the origin X must be in order that the magnitude of $\delta(x)$ be less than an arbitrarily small quantity, say δ. The answer is readily found to be $|x| > (\epsilon/\pi\delta)^{\frac{1}{2}}$. Similarly, if we ask ourselves how far from the origin X must be in order that $|\delta'(x)|$ be less than a small quantity, say δ', we find that we must have $|x| > (2\epsilon/\pi\delta')^{\frac{1}{3}}$. If, when the magnitudes of $\delta(x)$ and $\delta'(x)$ are respectively less than δ and δ', we may regard these magnitudes as being effectively equal to zero, we may write :

$$\delta(x) = 0 \quad \text{for} \quad |x| > \left(\frac{\epsilon}{\pi\delta}\right)^{\frac{1}{2}} \tag{1.12}$$

$$\delta'(x) = 0 \quad \text{for} \quad |x| > \left(\frac{2\epsilon}{\pi\delta'}\right)^{\frac{1}{3}} \tag{1.13}$$

and, in general, for the n th derivative of the delta function

$$\delta^{(n)}(x) = 0 \quad \text{for} \quad |x| > \left[\frac{(n+1)!\ \epsilon}{\pi\,\delta^{(n)}}\right]^{\frac{1}{n+2}} \tag{1.14}$$

Condition (1.12) is now to replace (1.2).

It will now be seen that for sufficiently small ϵ, namely for $\epsilon < 4\pi \frac{\delta^3}{\delta'^2}$, condition (1.12) can be satisfied

without (1.13) being at the same time satisfied. This means that $\delta'(x)$ is more "spread out" around the origin than $\delta(x)$ is. And, in general, the derivatives $\delta'(x)$, $\delta''(x)$, $\delta'''(x)$,... etc. become successively more spread out.

The effect of the derivatives of the delta function on an integrand may be deduced by a formal procedure of integration by parts. We find, for $a > 0$,

$$\int_{-a}^{a} f(x)\,\delta'(x)\,dx = \Big[f(x)\,\delta(x)\Big]_{-a}^{+a} - \int_{-a}^{+a} f'(x)\,\delta(x)\,dx$$

$$= -f'(0) \tag{1.15}$$

and, in general,

$$\int_{-a}^{+a} f(x)\,\delta^{(n)}(x)\,dx = (-1)^n f^{(n)}(0) \tag{1.16}$$

In equation (1.16) the function $f(x)$ is completely arbitrary, except that it must possess an nth derivative at the origin. In the theory of L. Schwartz the successive derivatives of the delta function are introduced through the concept of "distributions of increasing rank."

From (1.16) we may immediatly infer

$$\int_{-a}^{+a} f(x)\,x^m\,\delta^{(n)}(x)\,dx = (-1)^n \left[\frac{d^n}{dx^n}\left[x^m f(x)\right]\right]_{x=0}$$

$$= \begin{cases} (-1)^n \dfrac{n!}{(n-m)!}\, f^{(n-m)}(0) & \text{for } 0 \leq m \leq n \\ 0 & \text{for } m > n \end{cases} \tag{1.17}$$

which implies

$$x^m \delta^{(n)}(x) = \begin{cases} (-1)^m \dfrac{n!}{(n-m)!} \delta^{(n-m)}(x) & \text{for } 0 \leq m \leq n \\ 0 & \text{for } m > n \end{cases} \tag{1.18}$$

In particular

$$x \delta'(x) = -\delta(x) \tag{1.19}$$

$$x^2 \delta'(x) = 0 \tag{1.20}$$

$$x \delta''(x) = -2 \delta'(x) \tag{1.21}$$

Considerable care should be exercised in the application of these formal equations. We shall meet later an important case in which equation (1.21) is seemingly violated. At this point we can merely say that if the above expressions appear directly in integrals where the variable of integration is x, then these relations are generally valid. If, however, x is a function of other variables which are themselves the variables of integration, then these relations are only sometimes valid. This is not surprising, since the limiting process has so far been carried out only for a single well-defined path of integration (i.e. the real axis) with respect to a single variable x.

We now deduce a more or less miscellaneous set of relations involving the delta function, which will be of use later on.

Since

$$\int_{-a}^{a} g(x) f(x) \delta(x) dx = g(0) f(0)$$

we may infer

$$f(x) \delta(x) = f(0) \delta(x) \tag{1.22}$$

In particular

$$x\,\delta(x) = 0 \tag{1.23}$$

Since, for $\alpha \neq 0$,

$$\int_{-a}^{+a} f(x)\,\delta(\alpha x)\,dx = \frac{1}{\alpha}\int_{-\alpha a}^{\alpha a} f\left(\frac{y}{\alpha}\right)\delta(y)\,dy \qquad (y=\alpha x)$$

$$= \frac{1}{\alpha}\,\frac{\alpha}{|\alpha|}\,f(0) = \frac{1}{|\alpha|}\,f(0)$$

we infer

$$\delta(\alpha x) = \frac{1}{|\alpha|}\,\delta(x) \tag{1.24}$$

Since, for $a > 0$,

$$\int_{-\infty}^{+\infty} f(x)\,\delta(x^2-a^2)\,dx = \frac{1}{2}\int_{\infty}^{-a^2} f\left(-\sqrt{y+a^2}\right)\delta(y)\,\frac{dy}{\sqrt{y+a^2}} \qquad (y = x^2+a^2)$$

$$+ \frac{1}{2}\int_{-a^2}^{\infty} f\left(\sqrt{y+a^2}\right)\delta(y)\,\frac{dy}{\sqrt{y+a^2}}$$

$$= \frac{1}{2a}\left[f(-a) + f(a)\right]$$

we infer

$$\delta(x^2-a^2) = \frac{1}{2a}\left[\delta(x+a) + \delta(x-a)\right] \tag{1.25}$$

Differentiating equation (1.5), using (1.7), and remembering the implicit limiting convention, we obtain

$$\frac{d}{dx}\,\frac{x}{|x|} = 2\,\delta(x) \tag{1.26}$$

But

$$\frac{x}{|x|} = \frac{d}{dx}|x| \tag{1.27}$$

Hence $$\frac{d^2}{dx^2}|x| = 2\delta(x) \tag{1.28}$$

It is evident from the structure of the delta function that its integral is a "step function". That is,

$$\int_{-\infty}^{x} \delta(x)\,dx = \frac{1}{2}\left(1 + \frac{x}{|x|}\right) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x > 1 \end{cases} \tag{1.29}$$

Equation (1.29) checks with (1.26)

If the limiting process is taken quite seriously we may write

$$\left[\frac{x}{|x|}\right]_{x=0} = \lim_{\epsilon \to 0}\left[\frac{2}{\pi}\tan^{-1}\frac{x}{\epsilon}\right]_{x=0} = \lim_{\epsilon \to 0}[0] = 0 \tag{1.30}$$

and

$$\int_{-\infty}^{0} \delta(x)\,dx = \frac{1}{2}\,, \quad \int_{0}^{\infty} \delta(x)\,dx = \frac{1}{2} \tag{1.31}$$

Of importance in connection with the delta function is the logarithmic function, $\log x$. Since $\log x$ has an infinite number of branches at the origin and is uniquely determined only up to multiples of $2\pi i$, we here restrict our attention to two important forms, called $\log_+ x$ and $\log_- x$, along the real axis. These are defined as follows

$$\begin{aligned} \log_+ x &= \operatorname{Log}|x| + \frac{1}{2}\left(1 - \frac{x}{|x|}\right)\pi i \\ \log_- x &= \operatorname{Log}|x| - \frac{1}{2}\left(1 - \frac{x}{|x|}\right)\pi i \end{aligned} \tag{1.32}$$

where $\operatorname{Log}|x|$ denotes the real logarithm. Two important functions,

$\delta_+(x)$ and $\delta_-(x)$, are defined in terms of the derivatives of $\log_+(x)$ and $\log_-(x)$:

$$\left.\begin{aligned} -2\pi i\,\delta_+(x) &= \frac{d}{dx}\log_+(x) = \frac{1}{|x|}\,\frac{x}{|x|} - \pi i\,\delta(x) \\ 2\pi i\,\delta_-(x) &= \frac{d}{dx}\log_-(x) = \frac{1}{|x|}\,\frac{x}{|x|} + \pi i\,\delta(x) \end{aligned}\right\} \quad (1.33)$$

where use has been made of Eqs. (1.26) and (1.27) in taking the derivatives. Since $\frac{1}{|x|}\,\frac{x}{|x|} = \frac{1}{x}$, we have

$$\left.\begin{aligned} \delta_+(x) &= \frac{1}{2}\,\delta(x) + \frac{i}{2\pi}\,\frac{1}{x} \\ \delta_-(x) &= \frac{1}{2}\,\delta(x) - \frac{i}{2\pi}\,\frac{1}{x} \end{aligned}\right\} \quad (1.34)$$

Evidently

$$\delta(x) = \delta_+(x) + \delta_-(x) \quad (135)$$

Delta functions in Euclidean spaces of an arbitrary number n of dimensions can be readily defined in terms of the 1-dimensional delta function. For example, if $X_1, X_2, \ldots X_n$ are the coordinates of a point X in such a space, the n-dimensional delta function is

$$\delta(x) = \delta(x_1)\,\delta(x_2)\cdots\delta(x_n) \quad (1.36)$$

and satisfies

$$\int_R f(x)\,\delta(x)\,d_n x = f(0) \quad \text{for all } f(x) \quad (1.37)$$

where R is any region containing the origin and the symbol $d_n x$ denotes the n-dimensional volume element :

$$d_n x \equiv dx_1\,dx_2\cdots dx_n \quad (1.38)$$

The arbitrariness in R and $f(x)$ allows us to write formally

$$\delta(x) = 0 \qquad \text{for } x \neq 0 \tag{1.39}$$

$$\int_{\infty} \delta(x)\, d_n x = 1, \tag{1.40}$$

in analogy with (1.2) and (1.3).

Let us consider for a moment the important case of 3 dimensions. It is customary to label the points in 3-dimensional Euclidean space by the radius vector r.

$$r = (x_1, x_2, x_3) \tag{1.41}$$

Consider the function $1/\eta$ where $\eta \equiv |r| \equiv \sqrt{x_i x_i}$
For $r \neq 0$ we have, using customary notation of vector analysis,

$$\nabla \frac{1}{\eta} = -\frac{r}{\eta^3} \tag{1.42}$$

and $$\nabla^2 \frac{1}{\eta} = -\nabla \cdot \frac{r}{\eta^3} = -\frac{3}{\eta^3} + 3\, r \cdot \frac{r}{\eta^5} = 0 \tag{1.43}$$

But, by Gauss's theorem

$$\int_{\infty} \nabla^2 \frac{1}{\eta}\, d_3 r = -\int_{\infty} \nabla \cdot \frac{r}{\eta^3}\, d_3 r$$

$$= -\lim_{R \to \infty} \int_S \frac{1}{R^2}\, d\sigma = -4\pi \tag{1.44}$$

where $d\sigma$ is a surface element on the sphere S of radius R centered at the origin. Evidently we may write

$$\delta(r) = -\frac{1}{4\pi} \nabla^2 \frac{1}{\eta} \tag{1.45}$$

The function $-\frac{1}{4\pi}\frac{1}{\eta}$ is said to be the Green's function of the Laplacian operator ∇^2.

In general, The Green's function of any linear operator in n-dimensional Euclidean space is formally defined by

$$G_A(x) = A^{-1}\delta(x) \tag{1.46}$$

and satisfies

$$A\,G_A(x) = \delta(x) \tag{1.47}$$

If $\varphi(x)$ is a function satisfying an equation of the form

$$A\varphi = f \tag{1.48}$$

then a particular solution for φ is given by

$$\varphi(x) = \int G_A(x-x')\,f(x')\,d_n x' \tag{1.49}$$

From Eq. (1.28) it is seen that the Green's function for the operator $\frac{d^2}{dx^2}$ in 1 dimension is simply $\frac{1}{2}|x|$. For a 4-dimensional Euclidean space the Green's function of the operator

$$\Box^2 \equiv \frac{\partial^2}{\partial x_\mu \partial x_\mu} \quad , (x_\mu) = (x_1, x_2, x_3, x_4) \tag{1.50}$$

is readily found to be

$$G_{\Box^2}(x) = -\frac{1}{4\pi^2}\,\frac{1}{x_\mu x_\mu} \tag{1.51}$$

Use is made here of the fact that the surface volume of a 4-dimensional hypersphere of radius R is $2\pi^2 R^3$.

Equation (1.51) is valid only if the 4-dimensional space is positive definite. In the much more interesting case in which the space possesses a Minkowski metric, corresponding to the space-time of relativity theory, the Green's function of the d'Alembertian operator $\Box^2$ looks quite different from (1.51). In order to find out what it actually is we must first review the customary notation of relativity theory.

Of the coordinates x_1, x_2, x_3, x_4 the first three are real and the fourth is pure imaginary. The connection with non-relativistic notation is via

$$x_\mu = (x_i, x_4) = (r, ict) \tag{1.52}$$

where Latin indices range over the values 1,2,3, C is the velocity of light, and t is the time. The invariant volume element is

$$d_4 x = dx_1\, dx_2\, dx_3\, dx_0 \tag{1.53}$$

where $$x_0 = -i x_4 = ct \tag{1.54}$$

and the delta function is given by

$$\delta(x) = \delta(x_1)\,\delta(x_2)\,\delta(x_3)\,\delta(x_0) \tag{1.55}$$

Consider now the function $\delta(x_\mu x_\mu)$. Writing $x_\mu x_\mu = x^2$, we have

$$\frac{\partial}{\partial x_\mu}\delta(x^2) = 2x_\mu \delta'(x^2), \tag{1.56}$$

$$\Box^2 \delta(x^2) = 8\delta'(x^2) + 4x^2\delta''(x^2) \tag{1.57}$$

Referring to equation (1.21), we should be tempted to say $\Box^2\delta(r^2) = 0$. We shall see, however, that this is a case in which equation (1.21) is invalid. For, using (1.25), we may write

$$\delta(x^2) = \delta(-x^2) = \delta(x_0^2 - r^2) = \frac{1}{2r}\left[\delta(x_0+r) + \delta(x_0-r)\right],$$

and hence

$$\begin{aligned}
\Box^2 \delta(x^2) &= \left(\nabla^2 - \frac{\partial^2}{\partial x_0^2}\right)\left\{\frac{1}{2r}\left[\delta(x_0+r) + \delta(x_0-r)\right]\right. \\
&= -2\pi\,\delta(r)\left[\delta(x_0+r) - \delta'(x_0-r)\right] \\
&\quad - \frac{r}{r^3}\,\frac{r}{r}\left[\delta'(x_0+r) - \delta'(x_0-r)\right] \\
&\quad + \frac{1}{2r}\left\{\left(\frac{3}{r} - r\cdot\frac{r}{r^3}\right)\left[\delta'(x_0+r) - \delta'(x_0-r)\right]\right. \\
&\qquad \left.\left. + \frac{r}{r}\cdot\frac{r}{r}\left[\delta''(x_0+r) + \delta''(x_0-r)\right]\right\} \\
&\quad - \frac{1}{2r}\left[\delta''(x_0+r) + \delta''(x_0-r)\right] \\
&= -4\pi\,\delta(r)\,\delta(x_0) = -4\pi\,\delta(x)
\end{aligned} \tag{1.58}$$

Hence

$$G_{\Box^2}(x) = -\frac{1}{4\pi}\,\delta(x^2) \tag{1.59}$$

This Green's function is nowadays called the "invariant delta function" and denoted by $-\bar{D}(x)$

$$\bar{D}(x) \equiv \frac{1}{4\pi}\,\delta(x^2) \tag{1.60}$$

$$\Box^2\,\bar{D}(x) = -\delta(x) \tag{1.61}$$

It should be noted that the Green's function in the present case is actually not uniquely determined by (1.58). In place of (1.59) we could equally well take

$$G_{\Box^2}(x) = \frac{1}{4\pi}\,\delta(x^2) + f(x) \tag{1.62}$$

where $f(x)$ is any function which satisfies the free wave equation

$$\Box^2 f(x) = 0 \tag{1.63}$$

In a positive definite space, equation (1.63) would imply f = constant, and thus $G_{\Box^2}(x)$ would be uniquely determined except for an additive constant. In a Minkowski space, however, even the condition that $f(x)$ be an invariant function is not sufficient to restrict $G_{\Box^2}(x)$. As an example; we have $\Box^2 \frac{1}{x^2} = 0$, in contrast to the positive definite case. For, writing

$$\frac{1}{x^2} = \frac{1}{2x}\left[\frac{1}{x+x_0} + \frac{1}{x-x_0}\right]$$

and proceeding as in the case of $\delta(x^2)$, we have

$$\Box^2 \frac{1}{x^2} = -2\pi\,\delta(x)\left[\frac{1}{x+x_0} + \frac{1}{x-x_0}\right] = -2\pi\,\delta(x)\left[\frac{1}{x_0} - \frac{1}{x_0}\right] = 0 \tag{1.64}$$

The choice of an appropriate Green's function for a given problem will, in a Minkowski space, depend upon the boundary conditions for the problem.

A Green's function which has been found to be particularly useful in quantum field theory is the function $D_c(x)$, defined by

$$D_c(x) \equiv \frac{1}{\pi}\,\delta_+(x^2)$$

$$= \frac{1}{4\pi}\,\delta(x^2) + \frac{i}{4\pi^2}\,\frac{1}{x^2} \tag{1.65}$$

satisfying

$$\Box^2 D_c(x) = -\delta(x) \tag{1.66}$$

Introducing the function $D^{(1)}(x)$ defined by

$$D^{(1)}(x) = \frac{1}{2\pi^2}\,\frac{1}{x^2} \tag{1.67}$$

we may write

$$D_c(x) = \overline{D}(x) + \frac{1}{2}\,D^{(1)}(x) \tag{1.68}$$

2. Fourier series, integrals, and transforms.

Consider the following infinite series :

$$\sum_{n=-\infty}^{\infty} e^{(inx - |n|\epsilon)k}$$

$$= -1 + \sum_{n=0}^{\infty} e^{(ix-\epsilon)nk} + \sum_{n=0}^{\infty} e^{(-ix-\epsilon)nk}$$

$$= -1 + \frac{1}{1 - e^{(ix-\epsilon)k}} + \frac{1}{1 - e^{(-ix-\epsilon)k}}$$

$$= \frac{\sinh k\epsilon}{\cosh k\epsilon - \cos kx} \tag{2.1}$$

Let $f(x)$ be an arbitrary function. Then consider the following integral, for $a > 0$,

$$\sum_{n=-\infty}^{\infty} \int_{-a}^{a} f(x)\, e^{(inx - |n|\epsilon)k}\, dx$$

$$= \int_{-a}^{a} f(x) \frac{\sinh K\epsilon}{\cosh K\epsilon - \cos Kx} dx$$

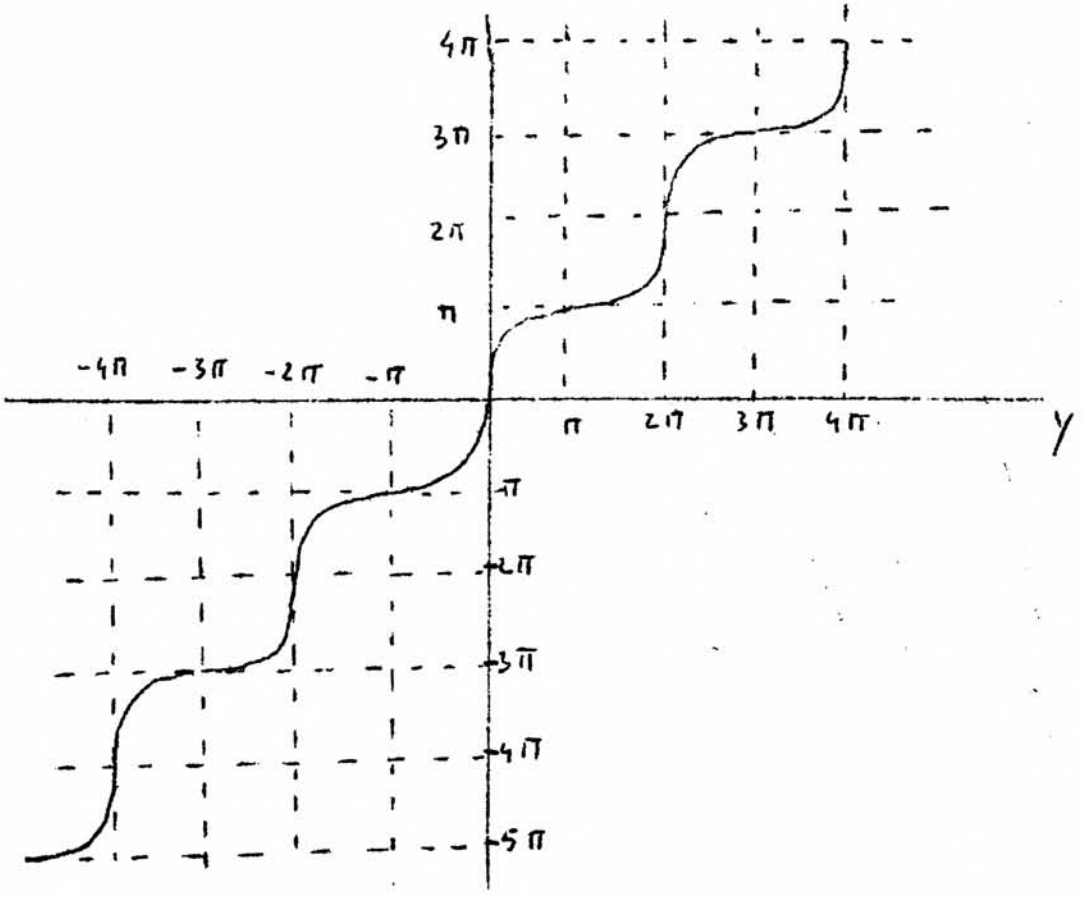
$$= \left[\frac{2}{K} f(x) \tan^{-1}\left(\sqrt{\frac{\cosh K\epsilon + 1}{\cosh K\epsilon - 1}} \tan \frac{Kx}{2} \right) \right]_{-a}^{+a}$$

$$- \frac{2}{K} \int_{-a}^{a} f'(x) \tan^{-1}\left(\sqrt{\frac{\cosh K\epsilon + 1}{\cosh K\epsilon - 1}} \tan \frac{Kx}{2} \right) dx \qquad (2.2)$$

We have here integrated by parts and used the elementary integral

$$\int \frac{\sqrt{a^2-1}}{a - \cos y} dy = 2 \tan^{-1}\left(\sqrt{\frac{a+1}{a-1}} \tan \frac{y}{2} \right) \qquad (2.3)$$

Now, the function $2 \tan^{-1}(A \tan \frac{y}{2})$ has the following appearance,



and it approaches more and more nearly to a "step-function" the larger larger A is. In particular, if $-2\pi < y < 2\pi$, then

$$\lim_{A\to\infty} \; 2\tan^{-1}\left(A \tan\frac{y}{2}\right) = \pi\,\frac{y}{|y|} \qquad (2.4)$$

Thus, taking the limit of equation (2.2) as $\epsilon \to 0$, and restricting a to the range $0 < a < \frac{2}{K}$, we obtain

$$\lim_{\epsilon\to 0} \sum_{n=-\infty}^{\infty} \int_{-a}^{a} f(x)\, e^{(inx-\ln|\epsilon|)k}\, dx$$

$$= \left[\frac{\pi}{k} f(x)\frac{x}{|x|}\right]_{-a}^{a} - \frac{\pi}{k}\int_{-a}^{a} f'(x)\frac{x}{|x|}\, dx = \frac{2\pi}{k} f(0) \qquad (2.5)$$

Evidently therefore the sum $\sum_{n=-\infty}^{+\infty} e^{(inx-\ln|\epsilon|)k}$ acts like a delta function. It is, however, unchanged if we add to x any multiple of $\frac{2\pi}{K}$. Hence it is rather a <u>sum of delta functions</u> equally spaced along the x axis. If we now extend the implicit limiting convention introduced in the previous section, by always supposing that an oscillating series of the form $\sum_{n=-\infty}^{+\infty} e^{inkx}$ is to be summed by inserting a small exponential convergence factor which is allowed to tend to zero only after all the analysis in which the series appears has been carried out, then we may write

$$\frac{K}{2\pi}\sum_{n=-\infty}^{\infty} e^{inKx} = \sum_{n=-\infty}^{\infty} \delta\left(x + \frac{2n\pi}{k}\right) \qquad (2.6)$$

Now, let $f(x)$ be any function which is periodic with period $\frac{2\pi}{K}$. (Mathematicians would impose restrictions of piecewise continuity.) Then we may write

$$f(x) = \int_{-\infty}^{+\infty} f(x')\,\delta(x-x')\,dx' = \int_{-\frac{\pi}{K}}^{\frac{\pi}{K}} \sum_{n=-\infty}^{\infty} f(x')\,\delta\left(x-x'+\frac{2n\pi}{K}\right)dx'$$

$$= \frac{K}{2\pi} \int_{-\frac{\pi}{K}}^{\frac{\pi}{K}} \sum_{n=-\infty}^{\infty} f(x') e^{ink(x-x')} dx' = \sum_{n=-\infty}^{\infty} f_n e^{inkx} \qquad (2.7)$$

where

$$f_n = \frac{K}{2\pi} \int_{-\frac{\pi}{K}}^{\frac{\pi}{K}} f(x) e^{-inkx} dx$$

$$= \frac{K}{2\pi} \int_{-\frac{\pi}{K}+a}^{\frac{\pi}{K}+a} f(x) e^{-inkx} dx \qquad (2.8)$$

The second expression above shows that , owing to the periodicity of $f(x)$, the domain of integration may be shifted forward or backward any distance a . Equations (2.7) and (2.8) express the fundamental relations of Fourier series.

Let us now see what happens when we take the constant K smaller and smaller. The period $\frac{2\pi}{K}$ becomes bigger and bigger, and in the limit it occupies the entire x-axis. We then have

$$\lim_{K\to 0} \frac{K}{2\pi} \sum_{n=-\infty}^{\infty} e^{inkx} = \lim_{K\to 0} \frac{K}{2\pi} \frac{\sin hk\epsilon}{\cosh k\epsilon - \cos kx}$$

$$= \frac{1}{\pi} \frac{\epsilon}{x^2+\epsilon^2} = \delta(x) \qquad (2.9)$$

The sum on the left may be replaced by an integral in the limit, through the correlation

$$nk \to k \qquad\qquad K \to dk \qquad (2.10)$$

Thus we have

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} dk = \delta(x) \qquad (2.11)$$

We shall presently verify this identification directly.

First let us evaluate two other oscillating integrals, by extending the implicit limiting convention to apply to integrals as well as series. The first of these is the following :

$$\frac{1}{2\pi}\int_0^\infty e^{ikx}\,dk = \frac{1}{2\pi}\int_0^\infty e^{(ix-\epsilon)k}\,dk$$

$$= -\frac{1}{2\pi}\,\frac{1}{ix-\epsilon} = \frac{i}{2\pi}\,\frac{x-i\epsilon}{x^2+\epsilon^2} = \frac{1}{2}\delta(x) + \frac{i}{2\pi}\,\frac{x}{x^2+\epsilon^2} \qquad (2.12)$$

The second is simply its complex conjugate :

$$\frac{1}{2\pi}\int_0^\infty e^{-ikx}\,dk = \frac{1}{2\pi}\int_{-\infty}^0 e^{ikx}\,dk = \frac{1}{2}\delta(x) - \frac{i}{2\pi}\,\frac{x}{x^2+\epsilon^2} \qquad (2.13)$$

The addition of (2.12) and (2.13) immediatly gives (2.11).

Now, in any integral in which the factor $\frac{x}{x^2+\epsilon^2}$ appears, the effect of passage to the limit $\epsilon \to 0$ is simply to take the "principal value" of the corresponding integral with $\frac{x}{x^2+\epsilon^2}$ replaced by $\frac{1}{x}$. This is often denoted symbolically by

$$\lim_{\epsilon\to 0}\,\frac{x}{x^2+\epsilon^2} = P\,\frac{1}{x} \qquad \text{(cf. Schwinger)} \qquad (2.14)$$

We shall, however, not adopt any special symbolism, but will leave it always understood that the principal value of an integral is to be taken whenever $1/x$ appears in it as a factor. Thus, remembering (1.34), we may write

$$\frac{1}{2\pi}\int_0^\infty e^{ikx}\,dk = \frac{1}{2}\delta(x) + \frac{i}{2\pi}\,\frac{1}{x} = \delta_+(x) \qquad (2.15)$$

$$\frac{1}{2\pi}\int_0^\infty e^{-ikx}\,dk = \frac{1}{2}\delta(x) - \frac{i}{2\pi}\,\frac{1}{x} = \delta_-(x) \qquad (2.16)$$

It is often convenient to write the functions $\delta_+(x)$ and $\delta_-(x)$ in the equivalent forms

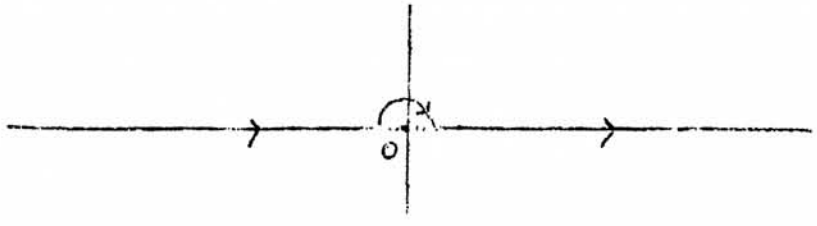
$$\delta_+(x) = \frac{i}{2\pi} \frac{1}{x + i\epsilon} \tag{2.17}$$

$$\delta_-(x) = \frac{-i}{2\pi} \frac{1}{x - i\epsilon} \tag{2.18}$$

The effect of the functions $\delta_+(x)$ and $\delta_-(x)$, occuring as factors in an integrand, can easily be demonstrated by making use of complex function theory. Let $f(x)$ be an arbitrary analytic function of Then consider the integral

$$\int_{-\infty}^{+\infty} f(x)\, \delta_+(x)\, dx = \frac{i}{2\pi} \int_{-\infty}^{+\infty} \frac{f(x)}{x + i\epsilon}\, dx \tag{2.19}$$

The integrand has a pole at $-i\epsilon$ which approaches the origin as ϵ goes to zero. The effect of passing to the limit $\epsilon \to 0$ is to indent the contour of integration in the following manner :



The residue of the integrand at the pole $-i\epsilon$ is

$$R_{-i\epsilon} = \frac{i}{2\pi} f(-i\epsilon) \longrightarrow \frac{i}{2\pi} f(0) \tag{2.20}$$

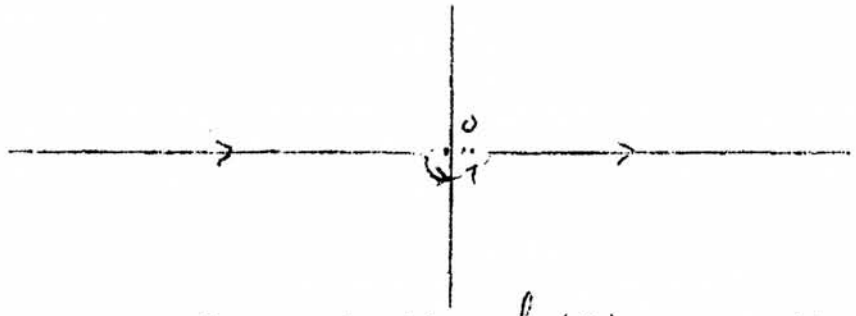
The contribution to the integrand arising from the indentation is

$$-\pi i\, R_{-i\epsilon} = \frac{1}{2} f(0)\,, \tag{2.21}$$

and the integral itself takes the value

$$\int_{-\infty}^{+\infty} f(x)\, \delta_+(x)\, dx = \frac{1}{2} f(0) + \frac{i}{2\pi} P \int_{-\infty}^{\infty} \frac{f(x)}{x}\, dx\,, \tag{2.22}$$

which cheks with (2.15). The function $\delta_-(x)$ can be discussed in a similar manner, the contour of integration taking in this case the form



For any function $f(x)$ we may write

$$f(x) = \int_{-\infty}^{+\infty} f(x')\,\delta(x-x')\,dx' = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x')\,e^{ik(x-x')}\,dk\,dx'$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f^{T*}(k)\,e^{ikx}\,dk \qquad (2.23)$$

where

$$f^{T}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f^{*}(x)\,e^{ikx}\,dx, \qquad (2.24)$$

provided that the integral (2.24) exists. The function f^T is known as the Fourier transform of the function f. Evidently

$$f^{TT}(x) \equiv f(x) \qquad (2.25)$$

That is, the process of taking the Fourier transform twice (or any even number of times) brings one back again to the original function. Equations (2.23) and (2.24) express the so-called Fourier integral theorem.

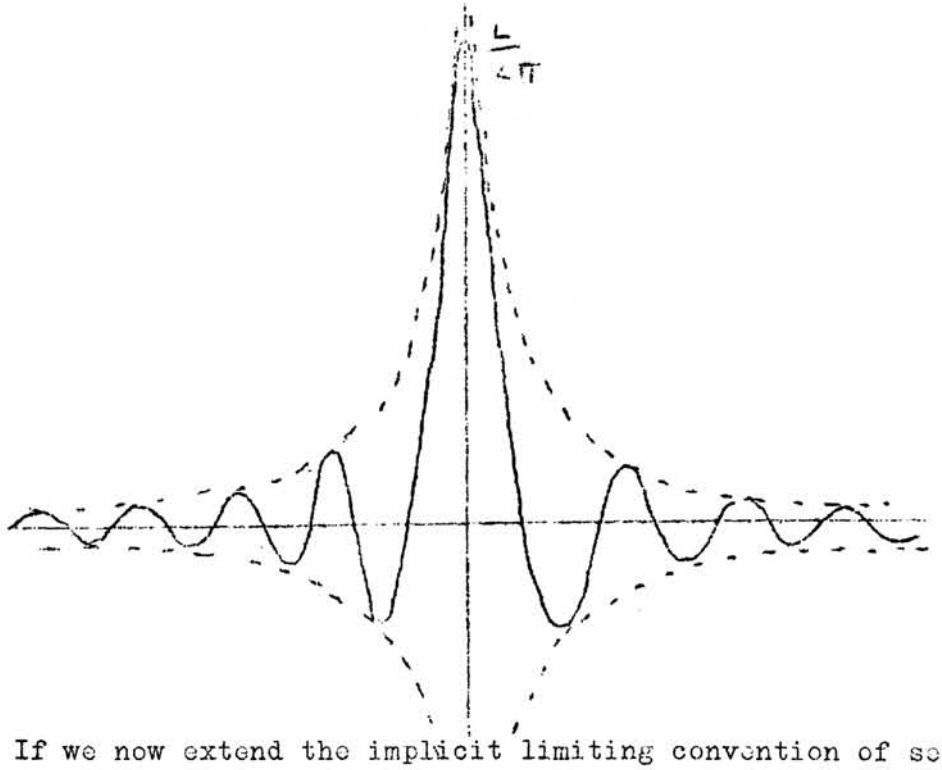
Consider now the integral

$$\frac{1}{2\pi}\int_{-\frac{L}{2}}^{\frac{L}{2}} e^{ikx}\,dx = \frac{1}{\pi}\,\frac{\sin k(L/2)}{k} \tag{2.26}$$

Referring to Eq. (2.11), we see that we may write formally

$$\lim_{L\to\infty}\ \frac{1}{\pi}\ \frac{\sin k\frac{L}{2}}{k} = \delta(k) \tag{2.27}$$

The function (2.26) has the following appearance



If we now extend the implicit limiting convention of section 1 to include L, so that the limit $L \to \infty$ is understood whenever L occurs in any expression, we may remove the sign $\lim_{L\to\infty}$ from the left hand side of (2.27). We may also write

$$\delta(0) = \frac{L}{2\pi} \to \infty \tag{2.28}$$

which is to be compared with (1.8). In order to make a distinction between k-space and x-space we shall subsequently always write

$\delta(0)$ in the form $\delta(k-k)$ or $\delta(x-x)$ according as we are working in one or the other of the two spaces. Thus Eq. (2.28) becomes

$$\bar{L} = 2\pi\,\delta(k-k) \tag{2.29}$$

Similarly, Eq. (1.8) of section 1 becomes

$$\frac{1}{\pi\epsilon} = \delta(x-x) \tag{2.30}$$

The quantity $\bar{L}$ appearing in (2.29) has the significance "lenght of the real axis". The cube of this quantity, $\bar{L}^3$, is thus to be understood as the "volume of space". We write

$$\bar{L}^3 = (2\pi)^3\,\delta(k-k) \tag{2.31}$$

The delta function appearing in (2.31) is the 3-dimensional delta function.

It is sometimes convenient to regard an integration over a certain region R of a space S formally as an explicit summation over the points of S lying in R. Thus we shall sometimes write

$$\int_R f(k)\,dk = \kappa \sum_{\substack{k \\ k \text{ in } R}} f(k) \tag{2.32}$$

where, in passing to the expression on the right, the symbol dk appearing on the left is thought of as a constant multiplier equal to an infinitesimally small quantity K (cf. (2.40)). The quantity K may readily be "evaluated" by introducing the Kronecker delta

$$\delta_{k,k'} = \begin{cases} 1 \text{ for } k = k' \\ 0 \text{ for } k \neq k' \end{cases} \tag{2.33}$$

From the properties (1.39, 40) of the delta function it is seen that we may write formally

$$\delta_{kk'} = \frac{\delta(k-k')}{\delta(k-k)} = \frac{\delta(k-k')}{\delta'(k'-k')} \tag{2.34}$$

Hence

$$k = k \sum_{k'} c_{kk'} = \int \frac{\delta(k-k')}{\delta(k-k)}\, dk' = \frac{1}{\delta(k-k)} \tag{2.35}$$

leading to

$$\int_R f(k)\, dk = \sum_{\substack{k \\ k \text{ in } R}} \frac{f(k)}{\delta(k-k)} = \left(\frac{2\pi}{L}\right)^n \sum_{\substack{k \\ k \text{ in } R}} f(k) \tag{2.36}$$

where n is the dimensionality of the space S.

We shall now calculate some miscellaneous Fourier transforms. First of all, we have the fundamental transform

$$\delta^T(k) = \frac{1}{\sqrt{2\pi}} \tag{2.37}$$

Next, we have, from (1.31),

$$\begin{aligned} \frac{x}{|x|} &= 2\int_0^x \delta(x)\, dx = \frac{1}{\pi}\int_0^x \int_{-\infty}^{+\infty} e^{ikx}\, dk\, dx \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \left[\frac{e^{ikx}}{ik}\right]_0^x dk = -\frac{1}{\pi i}\int_{-\infty}^{\infty} \frac{dk}{k} + \frac{1}{\pi i}\int_{-\infty}^{\infty} \frac{e^{ikx}}{k}\, dk \\ &= -\frac{i}{\pi}\int_{-\infty}^{\infty} \frac{e^{ikx}}{k}\, dk \end{aligned} \tag{2.38}$$

Hence

$$\frac{1}{k} = -\frac{i}{2}\int_{-\infty}^{\infty} \frac{x}{|x|} e^{ikx}\, dx \tag{2.39}$$

Integration or differentiation of these equations evidently leads to

other Fourier integral representations.

Differentiating the delta function we obtain the representation

$$\delta'(x) = \frac{i}{2\pi} \int_{-\infty}^{\infty} k\, e^{ikx}\, dk \tag{2.40}$$

which implies

$$\delta'^{T}(k) = -\frac{i}{\sqrt{2\pi}}\, k \tag{2.41}$$

And, in general, for the nth derivative,

$$\delta^{(n)T}(k) = \frac{1}{\sqrt{2\pi}}\, (-ik)^n \tag{2.42}$$

Fourier transforms in three dimensions are also frequently of importance. Among these we may mention the Fourier transform of the Green's function of the Laplacian operator. Referring to eq. (1.45), we may write

$$\frac{1}{r} = -4\pi \nabla^{-2} \delta(r) = -\frac{1}{2\pi^2} \nabla^{-2} \int_{\infty} e^{ik\cdot r}\, d_3 k$$

$$= \frac{1}{2\pi^2} \int_{\infty} \frac{e^{ik\cdot r}}{k^2}\, d_3 k \tag{2.43}$$

from which we obtain the inverse

$$\frac{1}{k^2} = \frac{1}{4\pi} \int_{\infty} \frac{e^{ik\cdot r}}{r}\, d_3 r \tag{2.44}$$

Equation (2.43) may be verified directly, using an infinitesimal exponential convergence factor. To show this, however, we shall consider a slightly more general case. We shall find the Green's function of the operator $\nabla^2 - \kappa^2$. We have

$$(\nabla^2 - \kappa^2)^{-1} \delta(r) = -\frac{1}{8\pi^3} \int_{\infty} \frac{e^{ik\cdot r}}{k^2 + \kappa^2}\, d_3 k \tag{2.45}$$

But

$$\int_\infty \frac{e^{-kr}}{r} e^{ik\cdot r} d_3 r = 2\pi \int_{-1}^{+1} dx \int_0^\infty r^2 dr \; \frac{e^{-kr}}{r} e^{i|k|rx}$$

$$= \frac{4\pi}{|k|} \int_0^\infty e^{-kr} \sin |k| r \, dr = \frac{4\pi}{k^2 + k^2} \qquad (2.46)$$

and hence

$$(\nabla^2 - k^2)^{-1} \delta(r) = -\frac{1}{4\pi} \frac{e^{-kr}}{r} \qquad (2.47)$$

Expression (2.47) is known as the Yukawa function.

One can see from eq. (2.45) how easy it is to write down the Fourier transform of the Green's function of a given differentiel operator. Once the Fourier transform has been found, the actual calculation of the function is simply a matter of straightforward evaluation of definite integrals, and the determination of Green's functions is thus reduced to a systematic procedure. The method can be very successfully employed to calculate various four dimensional Green's functions, of particular importance being those assiociated with the Minkowski space of special relativity, of which we have already met a few examples ((1.59), (1.65)). The study of these four dimensional Green's functions will, however, be deferred until the next chapter. In the meantime, we shall obtain a few further miscellaneous results.

Consider the integral

$$I = \int_{-\infty}^{\infty} e^{iax^2} dx \qquad (2.48)$$

where a is a complex number with $\Re a > 0$. We may write

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ia(x^2+y^2)} dx\, dy = 2\pi \int_0^\infty e^{iar^2} r\, dr$$

$$= \Pi \int_0^\infty e^{iax} dx = \frac{\Pi i}{a} = \frac{\Pi}{|a|} e^{\left(\frac{\Pi}{2} - \varphi\right) i} \tag{2.49}$$

where φ is the phase of a :

$$a = |a| e^{i\varphi} \tag{2.50}$$

Hence

$$I = \sqrt{\frac{\Pi}{|a|}} e^{\left(\frac{\Pi}{4} - \frac{\varphi}{2}\right) i} \tag{2.51}$$

By analytic continuation to the real axis we may write for $\varphi a = 0$, $a \neq 0$,

$$\int_{-\infty}^{\infty} e^{iax^2} dx = \sqrt{\frac{\Pi}{|a|}} e^{\frac{a}{|a|} \frac{\Pi}{4} i} , \tag{2.52}$$

Notions such as "fractional derivatives" or, more generally, the concept of the function of a differential operator may be introduced with the aid of Fourier transform theory. Thus we may write

$$f\left(\frac{d}{dx}\right) F(x) = \frac{1}{\sqrt{2\Pi}} \int_{-\infty}^{\infty} f(ik) \bar{F}^{T*}(k) e^{ikx} dk . \tag{2.53}$$

We have already used this formal procedure in the calculation of Green's functions. It is also a powerful method for generating useful approximations to the delta function. For example, if f is an arbitrary function, we may write

$$\delta(x) = \lim_{\epsilon \to 0} e^{i\epsilon f\left(\frac{d}{dx}\right)} \delta(x)$$

$$= \lim_{\epsilon \to 0} \frac{1}{2\Pi} \int_{-\infty}^{\infty} e^{i\left[\epsilon f(ik) + kx\right]} dk \tag{2.54}$$

The case in which $f(ik) = i|k|$ we have already met, and it leads to the expression (1.7) for the delta function. The case in which $f(ik) = -k^2$ is also of interest. We then have, using the implicit limiting convention $\epsilon \to 0$,

$$\delta(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i(-\epsilon k^2 + x k)} \, dk$$

$$= \frac{1}{2\pi} e^{i\frac{x^2}{4\epsilon}} \int_{-\infty}^{\infty} e^{-i\epsilon k'^2} \, dk' \qquad \left(k' = k - \frac{x}{2\epsilon}\right)$$

$$= \frac{e^{-\frac{\pi}{4}i}}{2\sqrt{\pi\epsilon}} e^{i\frac{x^2}{4\epsilon}} \qquad (2.55)$$

Another useful expression for the delta function is obtained for expressions (2.27) and (2.28). We may write formally

$$\delta(k) = \frac{\left[\delta(k)\right]^2}{\delta(k \cdot k)} = \frac{2}{L} \frac{\sin^2 k\frac{L}{2}}{\pi k^2} \qquad (2.56)$$

The expression on the right of (2.56) may readily be shown to have unit integral even when L is finite :

$$\int_{-\infty}^{\infty} \frac{2}{L} \frac{\sin^2 k\frac{L}{2}}{\pi k^2} \, dk = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} \, dx$$

$$= \frac{1}{\pi} \left\{ \left[-\frac{\sin^2 x}{x} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{\sin 2x}{x} \, dx \right\}$$

$$= \frac{1}{2\pi} \left[\frac{e^{ikx}}{ix} \right]_{k=-2}^{k=2} dx = \frac{1}{2\pi} \int dx \int_{-2}^{2} dk \, e^{ikx}$$

$$= \int_{-2}^{+2} \delta(k)\, dk = 1 \qquad (2.57)$$

Moreover,

$$\lim_{L_1 \to \infty} \frac{2}{L} \frac{\sin^2 k \frac{L}{2}}{\pi k^2} = \begin{cases} \infty & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} \qquad (2.58)$$

3. Space-like surfaces and the invariant Green's functions.

In this chapter we use the usual notation $(x_\mu) = (r, ict)$ to label the coordinates of space-time. Differentiation with respect to a coordinate will be denoted by a comma followed by an index.

A space-like surface σ may always be taken as one of a family of surfaces

$$\varphi = \text{constant} \qquad (3.1)$$

where φ is a real function whose value increases in positive time-like directions. The unit normal vector to σ at any point may then be defined by

$$\eta_\mu = \varphi_{,\mu} \left(-\varphi_{,\nu}\, \varphi_{,\nu}\right)^{-\frac{1}{2}} , \quad \eta_\mu\, \eta_\mu = -1 \qquad (3.2)$$

and can always be transformed by a proper Lorentz transformation into the form

$$(\eta_\mu) = (0, 0, 0, -i) \qquad (3.3)$$

Normal and tangential derivatives to the space-like surface are defined respectively by

$$\frac{\partial}{\partial n} \equiv -\eta_\mu \frac{\partial}{\partial x_\mu} , \qquad (3.4)$$

$$\bar{\nabla}_\mu \equiv \frac{\partial}{\partial x_\mu} - n_\mu \frac{\partial}{\partial \eta} \equiv \left(\delta_{\mu\nu} + n_\mu n_\nu \right) \frac{\partial}{\partial x_\nu} \qquad (3.5)$$

In particular, we have

$$\begin{aligned} \bar{\nabla}_\mu n_\nu &= \left[\varphi_{,\nu} \left(-\varphi_{,\alpha} \varphi_{,\alpha} \right)^{-\frac{1}{2}} \right]_{,\mu} \\ &\quad + \varphi_{,\mu} \left(-\varphi_{,\alpha} \varphi_{,\alpha} \right)^{-\frac{1}{2}} \varphi_{,\sigma} \left(-\varphi_{,\beta} \varphi_{,\beta} \right)^{-\frac{1}{2}} \left[\varphi_{,\nu} \left(-\varphi_{,\tau} \varphi_{,\tau} \right)^{-\frac{1}{2}} \right]_{,\sigma} \\ &= \varphi_{,\mu\nu} \left(-\varphi_{,\alpha} \varphi_{,\alpha} \right)^{-\frac{1}{2}} + \varphi_{,\nu} \varphi_{,\alpha} \varphi_{,\alpha\mu} \left(-\varphi_{,\beta} \varphi_{,\beta} \right)^{-\frac{3}{2}} \\ &\quad + \varphi_{,\mu} \varphi_{,\sigma} \varphi_{,\sigma\nu} \left(-\varphi_{,\beta} \varphi_{,\beta} \right)^{-\frac{3}{2}} + \varphi_{,\mu} \varphi_{,\nu} \varphi_{,\sigma} \varphi_{,\tau} \varphi_{,\sigma\tau} \left(-\varphi_{,\alpha} \varphi_{,\alpha} \right)^{-\frac{5}{2}} \end{aligned} \qquad (3.6)$$

Since the last expression is symmetric in μ and ν we have the theorem

$$\bar{\nabla}_\mu n_\nu = \bar{\nabla}_\nu n_\mu \qquad (3.7)$$

This theorem, involving only tangential derivatives, does not depend on the definition of σ as one of a family of surfaces. It depends only on the properties of the single surface itself (i.e. on how n_μ varies over σ) and is actually merely an expression of the fact that σ is <u>smooth</u> (i.e. differentiable).

Let $F[\sigma]$ be an arbitrary functional of σ . For example, $F[\sigma]$ may depend on the various values which certain fixed space-time quantities assume over σ , so that it depends in a definite way on the precise shape and location of σ . If χ is a point on σ then the variational derivative of $F[\sigma]$ with respect to σ at the point χ is defined by

$$\frac{\delta F[\sigma]}{\delta \sigma(\chi)} = \lim_{\Delta_4 \chi \to 0} \frac{F[\sigma'] - F[\sigma]}{\Delta_4 \chi} \qquad (3.8)$$

where σ' is a space-like surface which differs from σ only in the neighborhood of χ , and $\Delta_4 \chi$ is the 4-dimensional volume enclosed

between σ and σ'. For example, if

$$F[\sigma] = \int_{\sigma} f \, d\sigma_\mu \tag{3.9}$$

where $d\sigma$ is a surface element and

$$d\sigma_\mu = n_\mu \, d\sigma , \tag{3.10}$$

then, by the 4-dimensional Gauss theorem,

$$\frac{\delta F[\sigma]}{\delta \sigma(x)} = \lim_{\Delta_4 x \to 0} \left(\int_{\sigma'} - \int_{\sigma} \right) f \, d\sigma_\mu$$

$$= \lim_{\Delta_4 x \to 0} \frac{1}{\Delta_4 x} \int_{\Delta_4 x} f_{,\mu\mu} \, d_4 x = f_{,\mu\mu}(x) \tag{3.11}$$

Let f be any space-time function which vanishes faster than $1/r^2$ at sufficiently distant space-like directions. Then consider the surface integral

$$I_{\mu\nu}[\sigma] = \int_{\sigma} (n_\mu f_{,\nu} - n_\nu f_{,\mu}) \, d\sigma \tag{3.12}$$

From (3.11) we get

$$\frac{\delta I_{\mu\nu}[\sigma]}{\delta \sigma[x]} = f_{,\nu\mu} - f_{,\mu\nu} = 0 \tag{3.13}$$

which means that $I_{\mu\nu}$ is independent of σ, and to evaluate it we may choose σ as we please. If we choose σ to be a flat surface t = constant so that n_μ is given by (3.3), it is evident that the only non-vanishing components of $I_{\mu\nu}$ are $I_{i4} = -I_{4i}$ (i = 1, 2, 3). But

$$I_{i4} = -\int_{\sigma} n_4 f_{,i} \, d\sigma = i \int_{\infty} f_{,i} \, d_3 r = 0 \tag{3.14}$$

by the 3-dimensional Gauss theorem. Hence

$$\int_{\sigma} (\eta_\mu f_{,\nu} - \eta_\nu f_{,\mu})\, d\sigma = 0 \qquad (3.15)$$

From (3.5) it is seen that (3.15) may also be written in the form

$$\int_{\sigma} (\eta_\mu \bar{\nabla}_\nu - \eta_\nu \bar{\nabla}_\mu) f\, d\sigma = 0 \,, \qquad (3.16)$$

Using (3.16) we may evaluate the following integral

$$\int \bar{\nabla}_\mu f\, d\sigma = -\int_{\sigma} \eta_\nu \eta_\nu \bar{\nabla}_\mu f\, d\sigma = -\int_{\sigma} [\eta_\nu \bar{\nabla}_\mu (\eta_\nu f) - \eta_\nu f\, \bar{\nabla}_\mu \eta_\nu]\, d\sigma$$

$$= -\int_{\sigma} [\eta_\mu \bar{\nabla}_\nu (\eta_\nu f) - \eta_\nu f\, \bar{\nabla}_\nu \eta_\mu]\, d\sigma$$

$$= -\int_{\sigma} \eta_\mu (\bar{\nabla}_\nu \eta_\nu) f\, d\sigma = -\int k f\, d\sigma_\mu \qquad (3.17)$$

where

$$k = \bar{\nabla}_\mu \eta_\mu \qquad (3.18)$$

k is a measure of the curvature of σ . In the derivation of (3.17) we have also used the obvious identity

$$\eta_\mu \bar{\nabla}_\mu \equiv 0 \qquad (3.19)$$

In the relativistic theory of fields, of great importance is the function $\Delta_k(x)$ which satisfies the equation

$$(\Box^2 - k^2)\, \Delta_k = 0 \qquad (3.20)$$

and which is defined by the boundary conditions

$$\int_{\sigma_0} f\, \Delta_K \, d\sigma = 0 \tag{3.21}$$

and

$$\int_{\sigma_0} f\, \frac{\partial \Delta_K}{\partial n}\, d\sigma = -f(0) \tag{3.22}$$

where σ_0 is an arbitrary space-like surface through the origin and f is an arbitrary function over σ_0 . Equation (3.21) says that $\Delta_K(x)$ vanishes for $x^2 > 0$ as well as at the origin. Equation (3.22) says that the time derivative of $\Delta_K(x)$ is essentially a 3-dimensional delta function.

A number of properties of the function $\Delta_K(x)$ may be derived from the defining equations (3.20-22). Using the theorems derived previously, involving integrals over space-like surfaces, we have

$$\begin{aligned}\int_{\sigma_0} f \Delta_{K,\mu}\, d\sigma &= \int_{\sigma_0} f\left(\nabla_\mu \Delta_K + n_\mu \frac{\partial \Delta_K}{\partial n}\right) d\sigma \\ &= -n_\mu[0,\sigma_0] f(0) + \int_{\sigma_0} \left[\nabla_\mu (f\Delta_K) - \Delta_K \nabla_\mu f\right] d\sigma \\ &= -n_\mu[0,\sigma_0] f(0) - \int_{\sigma_0} (n_\mu k f - \nabla_\mu f)\, \Delta_K \, d\sigma \\ &= -n_\mu[0,\sigma_0] f(0)\end{aligned} \tag{3.23}$$

Here we denote the specific dependence of n_μ on σ_0 and on the point x at which it is to be evaluated on σ_0 by writing it $n_\mu[x,\sigma_0]$

An analogous integral involving the second derivatives of Δ_K can be evaluated in a similar fashion :

$$\int_{\sigma_0} f\, \Delta_{K,\mu\nu}\, d\sigma = \int_{\sigma_0} \left[(f\Delta_{K,\mu})_{,\nu} - f_{,\nu}\, \Delta_{K,\mu}\right] d\sigma$$

$$= n_\mu[0,\sigma_0] f_{,\nu}(0) + \int_{\sigma_0} \left[\nabla_\nu (f\Delta_{K,\mu})_{,\nu} - n_\nu n_\sigma (f\Delta_{K,\mu})_{,\sigma}\right] d\sigma$$

$$= n_\mu [\bar{0}, \sigma_0] f_{,\nu}(0)$$
$$+ \int_{\sigma_0} [\bar{} - n_\nu k f \Delta_{k,\mu} - n_\nu n_\sigma f_{,\sigma} \Delta_{k,\mu} - n_\nu n_\sigma f \Delta_{k,\mu\sigma}] d\sigma$$
$$= (n_\mu f_{,\nu} + n_\mu n_\nu k f + n_\mu n_\nu n_\sigma f_{,\sigma}) [\bar{0}, \sigma_0]$$
$$- \int_{\sigma_0} [\bar{} n_\sigma (n_\nu f \Delta_{k,\sigma})_{,\mu} - n_\sigma (n_\nu f)_{,\mu} \Delta_{k,\sigma}] d\sigma$$
$$= [\bar{} n_\mu n_\nu k f + n_\mu \bar{\nabla}_\nu f + (n_\nu f)_{,\mu}] [\bar{0}, \sigma_0] - \int_{\sigma_0} n_\mu (n_\nu f \Delta_{k,\sigma})_{,\sigma} d\sigma$$
$$= [\bar{} n_\mu n_\nu k f + n_\mu \bar{\nabla}_\nu f + (n_\nu f)_{,\mu} + n_\mu n_\sigma (n_\nu f)_{,\sigma}] [\bar{0}, \sigma_0]$$
$$- \int_{\sigma_0} n_\mu n_\nu f \Delta_{k,\sigma\sigma} d\sigma$$
$$= [\bar{} n_\mu n_\nu k f + n_\mu \bar{\nabla}_\nu f + \bar{\nabla}_\mu (n_\nu f)] [\bar{0}, \sigma_0]$$
$$- K^2 \int_{\sigma_0} n_\mu n_\nu f \Delta_k d\sigma$$
$$= [\bar{} f (n_\mu n_\nu k + \bar{\nabla}_\mu n_\nu) + (n_\mu \bar{\nabla}_\nu + n_\nu \bar{\nabla}_\mu) f] [\bar{0}, \sigma_0] \qquad (3.24)$$

Although, in the derivation of this result, we have used total derivatives of the normals (as in expressions of the form $(n_\nu f)_{,\mu}$ and have thereby tacitly assumed the possibility of expressing σ_0 as a member of a one-parameter family, ψ = constant, the final result involves only tangential derivatives and therefore depends only on the properties of the simple surface σ_0 and not on the particular choice for the rest of the family ψ = constant. It will be observed that the final expression in (3.24) is symmetric in μ and ν as, of course, it must be.

Since $\Delta_{k,\mu}$ has the nature of a 3-dimensional delta function, $\Delta_{k,\mu\nu}$ is the derivative of such a function. In the same limiting sense in which we wrote down equations (1.12,13) we

may also write

$$\Delta_{K,\mu} = 0 \quad \text{for} \quad x^2 > \frac{\epsilon}{\pi\sigma}, \tag{3.25}$$

$$\Delta_{K,\mu\nu} = 0 \quad \text{for} \quad x^2 > \left(\frac{2\epsilon}{\pi\sigma}\right)^{\frac{2}{3}} \tag{32.6}$$

indicating that $\Delta_{K,\mu\nu}$ is more "spread out" in space-like directions than $\Delta_{K,\mu}$ is. If x and x' are any two points on a space-like surface, then we shall take the expression $x \neq x'$ to be synonymous with $(x - x')^2 > \epsilon/\pi\sigma$. Thus, although $x \neq x'$ implies $\Delta_{K,\mu}(x-x') = 0$, it does not necessarily imply that $\Delta_{K,\mu\nu}(x-x')$ vanishes.

Consider the expression $n_\mu[\bar{x}';\sigma]\, n_\nu[\bar{x}';\sigma]\, f(x')$ i.e. the product of two components of the unit normal vector to an arbitrary space-like surface σ at the point x' on σ with an arbitrary function $f(x')$. This expression is a functional of σ, and it is sometimes useful to make use of the variational derivative of this functional. Writing

$$n_\mu[\bar{x}';\sigma]\, n_\nu[\bar{x}';\sigma]\, f(x') = -\int_\sigma f(x)\, \Delta_{K,\nu}(x-x')\, d\sigma_\mu$$

$$= -f(x') \int_\sigma \Delta_{K,\nu}(x-x')\, d\sigma_\mu \tag{3.27}$$

we may at once use (3.11) to calculate the variational derivative. Since one point of σ, namely x', is held fixed here, the variational derivative must be taken at a point x different from x'. Remembering, therefore, that $x \neq x'$ and hence $\Delta_{K,\nu}(x-x') = 0$ we have

$$\frac{\delta}{\delta\sigma(x)}\left\{n_\mu[\bar{x}';\sigma]\, n_\nu[\bar{x}';\sigma] f(x')\right\} = -f(x)\, \Delta_{K,\mu\nu}(x-x')$$

$$= -f(x')\, \Delta_{K,\mu\nu}(x-x') \tag{3.28}$$

Let f be an arbitrary solution of the equation

$$(\Box^2 - K^2) f = 0 \tag{3.29}$$

Consider the integral

$$I[x,\sigma] = \int_\sigma \left[f(x') \Delta_{k,\mu}(x'-x) - f_{,\mu}(x') \Delta_k(x'-x) \right] d\sigma'_\mu \tag{3.30}$$

where σ is an arbitrary space-like surface and x is an arbitrary space-time point which does not necessarily lie on σ. We have

$$\begin{aligned} \frac{\delta I[x,\sigma]}{\delta \sigma(x')} &= f_{,\mu}(x') \Delta_{k,\mu}(x'-x) + f(x') \Delta_{k,\mu\mu}(x'-x) \\ &\quad - f_{,\mu\mu}(x') \Delta_k(x'-x) - f_{,\mu}(x') \Delta_{k,\mu}(x'-x) \\ &= K^2 \left[f(x') \Delta_k(x'-x) - f(x') \Delta_k(x'-x) \right] = 0 \end{aligned} \tag{3.31}$$

Therefore I does not depend on σ and to evaluate it we may choose σ so as to pass through x. We immediatly obtain

$$f(x) = \int_\sigma \left[f(x') \Delta_{k,\mu}(x'-x) - f_{,\mu}(x') \Delta_k(x'-x) \right] d\sigma'_\mu \tag{3.32}$$

Equation (3.32) shows how the function Δ_k can be used to determine uniquely the effect of imposing boundary conditions on the solutions of equation (3.29). If the values of f and $\frac{df}{dn}$ are specified at all points of a space-like surface σ, then f is determined everywhere by (3.32).

In order to obtain an analytic expression for Δ_k it is convenient to introduce the following function :

$$\bar{\Delta}_k(x) = -\frac{1}{2} \varepsilon[x,\sigma_0] \Delta_k(x) \tag{3.33}$$

where the functional $\varepsilon[x, \sigma_0]$ is equal to + 1 or - 1 according as x lies to the future or the past of the arbitrary space-like surface σ_0 through the origin. Since $\Delta_K(x) = 0$ for $x^2 > 0$, the function $\bar{\Delta}_K$ does not depend on what surface σ_0 though the origin is chosen.

Let f be an arbitrary function which vanishes at distant space-time points. Consider the 4-dimensional integral

$$\int_\infty f \bar{\Delta}_{K,\mu\nu} \, d_4x = -\int_\infty f_{,\nu} \bar{\Delta}_{K,\mu} \, d_4x$$

$$= \frac{1}{2} \int_\infty f_{,\nu} \left(\varepsilon[x, \sigma_0] \Delta_K\right)_{,\mu} d_4x \qquad (3.34)$$

Now, $\left(\varepsilon[x, \sigma_0]\right)_{,\mu} \Delta_K(x) = 0$ because $\varepsilon[x, \sigma_0]$ is a step function and hence $\left(\varepsilon[x, \sigma_0]\right)_{,\mu}$ differs from 0 only on σ_0 where $\Delta_K(x)$ vanishes. We therefore have

$$\int_\infty f \bar{\Delta}_{K,\mu\nu} \, d_4x = \frac{1}{2} \int_\infty \varepsilon[x, \sigma_0] f_{,\nu} \Delta_{K,\mu} \, d_4x$$

$$= -\frac{1}{2} \int_{-\infty}^{\sigma_0} f_{,\nu} \Delta_{K,\mu} \, d_4x + \frac{1}{2} \int_{\sigma_0}^{\infty} f_{,\nu} \Delta_{K,\mu} \, d_4x$$

$$= \frac{1}{2} \int_{-\infty}^{\sigma_0} f \Delta_{K,\mu\nu} \, d_4x - \frac{1}{2} \int_{\sigma_0}^{\infty} f \Delta_{K,\mu\nu} \, d_4x$$

$$\cdots - \int_{\sigma_0} f \Delta_{K,\mu} \, d\sigma_\nu$$

$$= \eta_\mu[0, \sigma_0] \eta_\nu[0, \sigma_0] f(0) - \frac{1}{2} \int_\infty f \varepsilon[x, \sigma_0] \Delta_{K,\mu\nu} \, d_4x$$

$$= \int_\infty f(x) \left\{ -\frac{1}{2} \varepsilon[x, \sigma_0] \Delta_{K,\mu\nu}(x) + \eta_\mu[0, \sigma_0] \eta_\nu[0, \sigma_0] \delta(x) \right\} d_4x$$

(3.35)

Since f is arbitrary we have the theorem

$$\bar{\Delta}_{k,\mu\nu}(x) = -\frac{1}{2}\,\varepsilon[x,\sigma_0]\,\Delta_{k,\mu\nu}(x) + \eta_{\mu}[0,\sigma_0]\,\eta_{\nu}[0,\sigma_0]\,\delta(x) \tag{3.36}$$

As a corollary of (3.36) we find

$$(\Box^2 - k^2)\,\bar{\Delta}_k(x) = -\frac{1}{2}\,\varepsilon[x,\sigma_0]\,(\Box^2-k^2)\,\Delta_k(x) - \delta(x)$$

$$= -\delta(x) \tag{3.37}$$

$\bar{\Delta}_k$ is therefore evidently a Green's function of the operator $\Box^2 - k^2$.

In the case K = 0, $\bar{\Delta}_k$ becomes a Green's function of the d'Alembertian operator $\Box^2$. We have already studied the Green's functions of $\Box^2$ in chapter 3 of the text. Since $\bar{\Delta}_k(x) = 0$ for $x^2 > 0$, $\bar{\Delta}_0$ must be constructible as some combination of the Green's functions D^{ret} and D^{adv} (see eqs. (3.45) and (3.46) of the text) :

$$\bar{\Delta}_0(x) = a\,D^{ret}(x) + (1-a)\,D^{adv}(x)$$

$$= \frac{1}{4\pi}\left(1 + 2a\,\varepsilon[x,\sigma_0]\right)\delta(x^2)$$

$$= \bar{D}(x) + a\,D(x) - \left(\frac{2a-1}{2}\right)D(x)$$

(3.38)

Multiplying (3.38) by $-2\,\varepsilon[x,\sigma_0]$, we obtain

$$\Delta_0(x) = D(x) + 2a\,\bar{D}(x) - 2(2a-1)\,\bar{D}(x) \tag{3.39}$$

Applying the d'Alembertian operator to this equation and using eqs. (3.20) and (3.37) of this appendix and eq. (3.49) of the text, we find

$$0 = 4a\,\delta(x) - 2(2a-1)\,\delta(x) \tag{3.40}$$

which implies $2\alpha - 1 = 0$, giving

$$\bar{\Delta}_0(x) = \bar{D}(x) = \frac{1}{4\pi}\,\delta(x^2) \tag{3.41}$$

$$\Delta_0(x) = D(x) = -\frac{1}{2\pi}\,\varepsilon[x, c_0]\,\delta(x^2) \tag{3.42}$$

We may easily infer the form of $\bar{\Delta}_\kappa$ in the general case in which $\kappa \neq 0$. Since $\bar{\Delta}_\kappa(x) = 0$ for $x^2 > 0$, since $\bar{\Delta}_\kappa$ must reduce to $\bar{D}$ in the limit $\kappa \to 0$, and since $\bar{\Delta}_\kappa$ is an invariant function, no special coordinate system having been introduced in its definition, we write

$$\bar{\Delta}_\kappa(x) = \bar{D}(x) - \kappa^2\left(1 - \frac{\lambda}{|\lambda|}\right) f(\lambda) \tag{3.43}$$

where

$$\lambda = \kappa^2 x^2 \tag{3.44}$$

The form of the function $f(\lambda)$ is determined by the requirement that $\bar{\Delta}_\kappa$ satisfy eq. (3.37). Observing that $\lambda_{,\mu} = 2\kappa^2 x_\mu$, we have

$$\bar{\Delta}_{\kappa,\mu}(x) = \bar{D}_{,\mu}(x) + 4\kappa^4 x_\mu \delta(\lambda) f(\lambda) - 2\kappa^4 x_\mu\left(1 - \frac{\lambda}{|\lambda|}\right) f'(\lambda) \tag{3.45}$$

$$\bar{\Delta}_{\kappa,\mu\mu}(x) = -\delta(x) + 16\kappa^4 \delta(\lambda) f(\lambda) + 8\kappa^4 \lambda\,\delta'(\lambda) f(\lambda) + 8\kappa^4 \lambda\,\delta(\lambda) f'(\lambda)$$
$$- 8\kappa^4\left(1 - \frac{\lambda}{|\lambda|}\right) f'(\lambda) + 8\kappa^4 \delta(\lambda) f'(\lambda) - 4\kappa^4 \lambda\left(1 - \frac{\lambda}{|\lambda|}\right) f''(\lambda) \tag{3.46}$$

and hence, using (1.19, 22, 23, 24), we get

$$0 = \bar{\Delta}_{,\mu\mu}(x) - \kappa^2 \bar{\Delta}(x) + \delta(x)$$
$$= 8\kappa^2 f(0)\,\delta(x^2) - \frac{1}{4\pi}\,\kappa^2 \delta(x^2)$$
$$+ \kappa^4\left(1 - \frac{\lambda}{|\lambda|}\right)\left\{f(\lambda) - 8 f'(\lambda) - 4\lambda f''(\lambda)\right\} \tag{3.47}$$

which tells us that f is a solution of the differential equation

$$4\lambda \frac{d^2 f}{d\lambda^2} + 8 \frac{df}{d\lambda} - f = 0 \tag{3.48}$$

satisfying

$$f(0) = \frac{1}{32\pi} \tag{3.49}$$

Normally, the single condition (3.49) would not be sufficient to fix the solution of a second order differential equation such as (3.48). We shall see, however, that since the origin is a singular point of the differential equation (3.48), the condition (3.49) suffices in the present case. Let us first make the change of variables

$$u = (-\lambda)^{\frac{1}{2}} \qquad J = 16\pi u f \tag{3.50}$$

We have

$$\lambda = -u^2 \qquad \text{and} \qquad \frac{du}{d\lambda} = -\frac{1}{2}(-\lambda)^{-\frac{1}{2}} = -\frac{1}{2u} \tag{3.51}$$

so that

$$\frac{df}{d\lambda} = \frac{df}{du}\frac{du}{d\lambda} = -\frac{1}{2u}\frac{df}{du} \tag{3.52}$$

$$\frac{d^2 f}{d\lambda^2} = -\frac{1}{2u}\left(\frac{1}{2u^2}\frac{df}{du} - \frac{1}{2u}\frac{d^2 f}{du^2}\right) \tag{3.53}$$

and

$$0 = -4u^2\left(-\frac{1}{4u^3}\frac{df}{du} + \frac{1}{4u^2}\frac{d^2 f}{du^2}\right) - \frac{4}{u}\frac{df}{du} - f$$

$$= -\left(\frac{d^2 f}{du^2} + \frac{3}{u}\frac{df}{du} + f\right) \tag{3.54}$$

Also

$$\frac{df}{du} = \frac{1}{16\pi}\left[-\frac{J}{u^2} + \frac{1}{u}\frac{dJ}{du}\right] \tag{3.55}$$

and

$$\frac{d^2 f}{du^2} = \frac{1}{16\pi}\left[2\frac{J}{u^3} - \frac{2}{u^2}\frac{dJ}{du} + \frac{1}{u}\frac{d^2 J}{du^2}\right] \tag{3.56}$$

so that

$$0 = 2\frac{J}{u^3} - \frac{2}{u^2}\frac{dJ}{du} + \frac{1}{u}\frac{d^2 J}{du^2} - 3\frac{J}{u^3} + \frac{3}{u}\frac{dJ}{du} + \frac{J}{u}$$

$$= \frac{1}{u}\left[\frac{d^2 J}{du^2} + \frac{1}{u}\frac{dJ}{du} + \left(1 - \frac{1}{u^2}\right) J\right] \tag{3.57}$$

Equation (3.57) is Bessel's equation of order 1. From (3.49, 50) we see that the function $J(u)$ satisfies the condition

$$\lim_{u \to 0} \frac{J(u)}{u} = \frac{1}{2} \tag{3.58}$$

which identifies it as the Bessel function of the first kind of order 1

$$J(u) = J_1(u) = \frac{u}{2} - \frac{u^3}{2^2 \cdot 4} + \frac{u^5}{2^2 \cdot 4^2 \cdot 6} - \frac{u^7}{2^2 \cdot 4^2 \cdot 6^2 \cdot 8} + \cdots \tag{3.59}$$

Hence, finally,

$$\bar{\Delta}_K(x) = \frac{1}{4\pi}\delta(x^2) - \frac{K^2}{16\pi}\left(1 - \frac{x^2}{|x^2|}\right)\frac{J_1\left(K\sqrt{-x^2}\right)}{K\sqrt{-x^2}} \tag{3.60}$$

Since $\bar{\Delta}_K$ does not involve any preferred direction in space-time, as shown by the absence of any vectors η_μ, ϵ_μ, etc. in (3.61), its Fourier representation may be written down at once from (3.37) :

$$\bar{\Delta}_K(x) = -(\Box^2 - K^2)^{-1}\delta(x) = \frac{1}{(2\pi)^4}\int_\infty \frac{e^{ikx}}{k^2 + K^2}\, d_4 k \tag{3.61}$$

With the aid of the miscellaneous integrals evaluated in the preceding chapter this may be rewritten in the form

$$\bar{\Delta}_K(x) = -\frac{i}{2(2\pi)^4}\int_{-\infty}^{+\infty} du \int_\infty d_4 k\, \frac{a}{|a|}\, e^{i\left[(k^2 + K^2)a + kx\right]}$$

$$= -\frac{i}{2(2\pi)^4}\int_{-\infty}^{\infty} da \int_{\infty} d_4 k' \frac{a}{|a|}\, e^{i\left(K^2 a - \frac{x^2}{4a}\right)} e^{i a k'^2} \qquad \left(k' = k + \frac{x}{2a}\right)$$

$$= -\frac{i}{2(2\pi)^4}\int_{-\infty}^{\infty} \frac{a}{|a|}\, e^{i\left(K^2 a - \frac{x^2}{4a}\right)} \frac{\pi^2}{|a|^2}\, e^{3\frac{a}{|a|}\frac{\pi}{4} i}\, e^{-\frac{a}{|a|}\frac{\pi}{4} i}$$

$$= \frac{1}{32\pi^2}\int_{-\infty}^{\infty} \frac{1}{a^2}\, e^{i\left(K^2 a - \frac{x^2}{a}\right)} da$$

$$= \frac{1}{8\pi^2}\left(-\int_{0}^{-\infty} - \int_{+\infty}^{0}\right) e^{-i\left(x^2\alpha - \frac{K^2}{4\alpha}\right)} d\alpha$$

$$= \frac{1}{8\pi^2}\int_{-\infty}^{+\infty} e^{-i\left(x^2 d - \frac{K^2}{4d}\right)} dd \qquad (3.62)$$

where

$$\alpha = \frac{1}{4a} \qquad , d\alpha = -\frac{da}{4a^2} \qquad (363)$$

In the case K = 0, equation (3.62) is seen to reduce to immediatly to the form (3.41) :

$$\overline{D}(x) = \frac{1}{8\pi^2}\int_{-\infty}^{\infty} e^{-i x^2 \alpha} d\alpha = \frac{1}{4\pi}\,\delta(x^2) \qquad (3.64)$$

The Fourier representation of the function Δ_K is also readily obtained. Let ξ_μ be an arbitrary infinitesimal time-like vector with a negative time-like orientation. Then, using the first expression appearing in (3.62), we may write

$$\Delta_k(x) = -2\frac{\xi x}{|\xi x|}\,\overline{\Delta}_K(x) = \frac{2i}{\pi}\,\overline{\Delta}_K(x)\int_{-\infty}^{\infty} \frac{1}{z}\, e^{i\xi x z} dz$$

$$= \frac{1}{\pi(2\pi)^4}\int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} da \int d_4 k\, \frac{a}{|a|}\, e^{i\left[(k^2, K^2)a + kx + \xi x z\right]}$$

$$= \frac{1}{\pi(2\pi)^4}\int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} da \int_{\alpha} d_4 k'\, \frac{1}{z}\,\frac{a}{|a|}\, e^{i\left[k'^2 a - 2k'\xi z a + K^2 a + k'x\right]} \qquad (k' = k + \xi z)$$

$$\Delta_K(x) = -\frac{i}{(2\pi)^4} \int_{-\infty}^{\infty} da \int_{\infty} dk' \; \frac{k'\epsilon a}{|k'\epsilon a|} \; \frac{a}{|a|} \; e^{i\left[(k'^2+K^2)a + k'x\right]}$$

$$= -\frac{i}{2\pi^3} \int_{\infty} \frac{\epsilon k}{|\epsilon k|} \delta(k^2+K^2) e^{ikx} d_4 k \qquad (3.65)$$

Since $(k^2 + K^2)\,\delta(k^2 + K^2) = 0$, expression (3.65) shows directly that Δ_K satisfies the equation (3.20). In the case K = 0, equation (3.65) reduces to

$$D(x) = -\frac{1}{2\pi} \frac{\epsilon x}{|\epsilon x|} \delta(x^2) = \frac{i}{(2\pi)^2} \int_{\infty} D(k)\, e^{ikx} d_4 k \qquad (3.66)$$

Defining 4 -dimensional Fourier transforms by the equations

$$f(x) = \frac{1}{(2\pi)^2} \int_{\infty} f^T(k)\, e^{ikx} \; d_4 k \qquad (3.67)$$

$$f^T(k) = \frac{1}{(2\pi)^2} \int_{\infty} f(x)\, e^{ikx} d_4 x \,, \qquad (3.68)$$

we may evidently write

$$D^T = -iD \qquad (3.69)$$

That is, apart from a phase factor, the function D is equal to its own Fourier transform.

In general, any function f satisfying the equation $(\Box^2 - K^2) f = 0$ has a Fourier representation of the form

$$f(x) = \int_{\infty} g(k)\, \delta(k^2 + K^2)\, e^{ikx} d_4 k \qquad (3.70)$$

Since $k^2 + K^2$ vanishes only for vectors k_μ lying inside the light cone through the origin in k-space, the region of k-space for which the

integrand in (3.70) is non-vanishing and for which $k_0 > 0$ is completely separated from the region for which the integrand is non-vanishing and $k_0 < 0$. The function f may therefore be divided into positive and negative frequency parts as follows :

$$f^{(+)}(x) = \int g(k)\, \delta(k^2+K^2)\, e^{ikx}\, d_4 k, \tag{3.71}$$

$$f^{(-)}(x) = \int_{k_0<0}^{k_0>0} g(k)\, \delta(k^2+K^2)\, e^{ikx}\, d_4 k \tag{3.72}$$

The positive and negative frequency parts of the function Δ_K may be read off directly from (3.65). An important function occuring in the quantum theory of fields is the following :

$$\Delta_K^{(1)}(x) = i\left[\Delta_K^{(+)}(x) - \Delta_K^{(-)}(x)\right]$$

$$= \frac{1}{(2\pi)^3} \int_{\infty} \delta(k^2+K^2)\, e^{ikx}\, d_4 k \tag{3.73}$$

satisfying

$$(\Box^2 - K^2)\, \Delta_K^{(1)} = 0 \tag{3.74}$$

Its explicit form can be derived with the aid of the integrals of chapter 2 and the methods of (3.62) :

$$\Delta_K^{(1)}(x) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d\alpha \int_{\infty} d_4 k\, e^{i\left[(k^2+K^2)\alpha + kx\right]}$$

$$= \frac{i}{4\pi} \int_{-\infty}^{+\infty} \frac{\alpha}{|\alpha|}\, e^{-i\left(x^2\alpha - \frac{K^2}{4\alpha}\right)} d\alpha \tag{3.75}$$

In the case K = 0 this reduces to

$$D^{(1)}(x) = \frac{i}{4\pi^2} \int_{-\infty}^{\infty} \frac{\alpha}{|\alpha|}\, e^{-ix^2\alpha}\, d\alpha = \frac{1}{2\pi i} \frac{1}{x^2} \tag{3.76}$$

which is the same as (1.67).

Also important is the generalization of (1.68) to the case $K \neq 0$. We define

$$\Delta_{Kc}(x) = \overline{\Delta}_K(x) + \frac{i}{2}\Delta_K^{(1)}(x) \tag{3.77}$$

Using the integral representations (3.61) and (3.73), we have

$$\Delta_{Kc}(x) = \frac{i}{(2\pi)^3}\int_\infty \left[\frac{1}{2}\delta(k^2+K^2) - \frac{i}{2\pi}\frac{1}{k^2+K^2}\right] e^{ikx}\, d_4 k$$

$$= \frac{1}{(2\pi)^3}\int_\infty \Delta_c(k^2+K^2)\, e^{ikx}\, d_4 k \tag{3.78}$$

In the case $K = 0$ this reduces to

$$D_c(x) = \frac{1}{2\pi}\,\delta_+(x^2) = \frac{i}{(2\pi)^2}\int_\infty D_c^*(k^2)\, e^{ikx}\, d_4 k \tag{3.79}$$

from which we obtain

$$D_c^T = -i D_c \tag{3.80}$$

That is D_c, like D, is, apart from a phase factor, equal to its own Fourier transform.

For many purposes it is convenient to make use of the formal expression (2.18) and write (3.78) in the form

$$\Delta_{Kc}(x) = \frac{1}{(2\pi)^4}\int_\infty \frac{e^{ikx}}{k^2+K^2-i\epsilon}\, d_4 k \tag{3.81}.$$

Q U A N T U M M E C H A N I C S

Lecture Notes by

Bryce S. DE WITT
Radiation Laboratory
University of California

E R R A T A

Université de Grenoble
Cours professé à
L'ECOLE D'ETE DE PHYSIQUE THEORIQUE
Les Houches (Haute-Savoie) France
Juillet 1953

E R R A T A

Chapter 1.-

Page	Lines	Read :
1	1	n coordinates ρ^i
	11, 12	at least quadratic in the $\dot{\rho}^i$
6	15	subtracting
8	3	$\|A\| = \left[\frac{\partial(\rho,p)}{\partial(u)}\right]^2$
9	5	Hint : Write the expression
11	1	If we now carry out a canonical.....
	3, 4	$= \int_R \frac{1}{2^m} \ldots\ldots$
12	6, 7	and $\mathfrak{S}$ is an infinitesimal
14	9	two first order infinitesimal canonical....
16	7	two first order infinitesimal point.....
19	13	$\frac{\partial S}{\partial t} = -H, \quad \frac{\partial S}{\partial t'} = H'$
21	12	The first of Eqs. (1.87) may be combined
24	9	The constant may be adjusted....
	11	W is sometimes also called....
	12	conjugate to the time t,
	13	$S = \int_{\text{trajectory}} (p_i \, d\rho^i + \lambda \, dt)$
24	14	This suggests a simple distinction....
	16	If, in Eq. (1.98), p_1'' (or p_1''') is chosen....

27 — Insert a footnote reading "The notation NS following any equation indicates that "no summation" is to be carried out over repeated indices occuring in the equation."

14 — $$p_i = \frac{\partial W}{\partial q^i}, \quad NS. \qquad (1.45)$$

29 21 — ω^i remains unchanged over cycles j with $j \neq i$.

31 7 — ω changes by amount.

33 1 — If only one of the τ_i, say τ_1, is ....

21 — to the variables ω^A, J_A vanish, while those...

35 7 — express the ω's as functions of the q's and J's .

36 11 — Problem : For the harmonic oscillator show that

$$E = J\omega \qquad (1.185)$$

$$W = -\frac{m\omega}{2}\left[q\sqrt{\frac{2J}{m\omega} - q^2} - \frac{2J}{m\omega}\cos^{-1}\sqrt{\frac{m\omega}{2J}}\,q\right]$$

$$q = \sqrt{\frac{2J}{m\omega}}\cos\omega \qquad (1.190)$$

$$p = -\sqrt{2mJ\omega}\sin\omega$$

$$\omega = -\frac{m\omega}{2}q^2 \tan\omega = -\frac{1}{2}J\sin 2\omega$$

40 1

$$p = \frac{\partial W}{\partial q} = \frac{J}{\sqrt{1-q^2}} = \frac{J}{\cos\omega} \qquad (1.210)$$

42 4 $|\alpha_\varphi|$ must be less than α_θ and ....

43 12 $$J_r = \frac{1}{\pi}\int_{r_{min}}^{r_{max}} \sqrt{2m[E - V(r)] - \frac{(J_\theta + |J_\varphi|)^2}{r^2}}\, dr \qquad (1.231)$$

16 $$\omega_\theta = |\omega_\varphi| = \frac{\partial E}{\partial J_\theta} = \frac{\partial E}{\partial |J_\varphi|} \qquad (1.232)$$

44 9 $$\omega_m = \mp \omega_\theta + \omega_\varphi \qquad (1.238)$$

11 $$T = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & \mp 1 & 0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \pm 1 & \pm 1 & 1 \end{pmatrix} \qquad (1.239)$$

$\downarrow$

{ + sign when $J_\varphi \geq 0$
{ − sign when $J_\varphi \leq 0$

48 In lines 10 and 14 replace J_φ by $|J_\varphi|$.

Note

The Appendix is located at the end of chapter III (Vol. II).

Chapter II -

Page	Lines	Read :
2	19	... To zeroth order the transformation.....
6	16	... subtracting (2.22)...
8	21	infinitely more numerous than ....
11	16	$= \frac{iH_{1\tau}}{\pi\omega_0}$ for $\tau \neq 0$ (2.46)
12	5	$\omega^i = \frac{\partial E}{\partial J_i} = \frac{\partial E_0}{\partial J_i} + \frac{\partial E_1}{\partial J_i} + \ldots = \omega_0^i + \frac{\partial H_{10}}{\partial J_i} + \ldots$ (2.49)
13	5	$H_1(\omega_0, J_0, t) = H_1(\rho, t) = -L_1(\rho, t)$ (2.53)
	6	where L_1 is the perturbing term....
	7	The term L_1 describes....
	8	$F_1^{ext} = \frac{\partial L}{\partial \rho^i} = -\frac{\partial H_1}{\partial \rho^i}$ (2.54)

Chapter III.-

Front cover : For "Department of Physics", read : "Radiation Laboratory".

Page	Lines	Read :
1	7	$m_i \ddot{r}_i = \ldots$ (3.1)
3		The integration variable $d\pi$ should be replaced by $d\tau$ in Eqs. (3.3) and (3.5).

The symbol $\mathcal{R}$ should be replaced by r in lines 14 and 15.

8 21 the Green's functions of the d'Alembertian operator $\Box^2$.

9 9 $\Box^2+2\varepsilon\frac{\partial}{\partial x_\mu}$ can immediately.....

10 $$G_{\Box^2+2\varepsilon\frac{\partial}{\partial x}}(x)=-\frac{1}{(2\pi)^4}\int_\infty \frac{e^{ikx}}{k^2-2i\varepsilon k}\,d^4k \qquad (3.36)$$

19 $$G_{\Box^2+2\varepsilon\frac{\partial}{\partial x}}(x)=\ldots. \qquad (3.39)$$

In lines 7 and 19 the integration variable dk should be replaced by d^4k, and in line 20 it should be replaced by d^3k.

10 1, 2 $$\int_\infty \frac{1}{|k|}\left(e^{i|k|\alpha}-e^{-i|k|\alpha}\right)e^{ik\cdot r}\,d^3k$$
$$=2\pi\int_0^\infty\int_{-1}^{1}\frac{1}{|k|}\left(e^{i|k|\alpha}-e^{-i|k|\alpha}\right)|k|^2\,dx\,d|k|$$

7 $$G_{\Box^2+2\varepsilon\frac{\partial}{\partial x}}(x)=\ldots.$$

12 $$G_{\Box^2+2\varepsilon\frac{\partial}{\partial x}}(x)=\ldots.$$

11 3 $$G_{\Box^2+2\varepsilon\frac{\partial}{\partial x}}(x)=\ldots.$$

12 2 ... retarded potential of the current j_μ .

15 1, 2, 3

$$F^{adv}_{\mu\nu}(x) = A^{adv}_{\nu,\mu}(x) - A^{adv}_{\mu,\nu}(x)$$

$$= e\,\frac{\ddot{z}^a_\mu (z^a_\nu - x_\nu) - \ddot{z}^a_\nu (z^a_\mu - x_\mu)}{[\dot{z}^a (z^a - x)]^2}$$

$$- e\,\frac{[\ddot{z}^a_\mu (z^a_\nu - x_\nu) - \ddot{z}^a_\nu (z^a_\mu - x_\mu)][\dot{z}^a (z^a - x) + (\dot{z}^a)^2]}{[\dot{z}^a (z^a - x)]^3} \qquad (3.64)$$

24 16 $\bar{F}_{\mu\sigma} \bar{F}_{\nu\sigma}\, d\sigma_\nu = \ldots$

25 15 $$= -e F^{free}_{\mu\sigma} z_\sigma \, d\Omega \, d\tau \qquad (3.117)$$

28 In lines 19, 20 and 22, replace the symbol r by σ.

29 In lines 6, 7, 13, 15, 16, 17, 19, replace the symbol r by σ .

30 In lines 2, 3, replace r by σ .

22 $$\dot{r} = t\,\frac{dr}{dt} = \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (3.146)$$

31 2 $$ct = \sqrt{c^2 + \dot{r}^2}, \ldots$$

32 In lines 17, 19, 22 replace r by σ.

33 In lines 1, 3, 5, 10, 10, 11, 12, 14, 15 replace r by σ .

35	2	$= \frac{c}{4\pi} E^{in} \times H^{in}$
	8	(3.176')
	9	(3.176")
36		In lines 1, 5, 7 replace the symbols ρ and r by σ.
37		In lines 5, 7, 13, 19 replace ρ by σ.
38		In line 4 replace r by σ.

Appendix.

4	18	..... (see graph
21	5	..... Let f(x) be an arbitrary analytic...
24	20	Replace $\delta(k,k')$ by $\delta_{kk'}$.
25	1	Replace $\delta(k,k')$ by $\delta_{kk'}$.
	3	Replace $\delta(k,k')$ by $\delta_{kk'}$
26	8	$\delta^{(n)T}(k) = \frac{1}{\sqrt{2\pi}}(-ik)^n$ (2.42)
27	10	calculation of the function is.....
	22	where a is a complex number with $Im\, a > 0$.
35	20	Δ_k can be evaluated in a similar fashion:

44	2	$= -\frac{i}{2(2\pi)^4}\int_{-\infty}^{\infty}\frac{a}{\lvert a\rvert}e^{i(k^2a-\frac{z^2}{4a})}\frac{\pi^2}{\lvert a\rvert^2}e^{3\frac{a}{\lvert a\rvert}\frac{\pi}{4}i}e^{-\frac{a}{\lvert a\rvert}\frac{\pi}{4}i}da$
	15	$= \frac{1}{\pi(2\pi)^4}\int_{-\infty}^{\infty}dz\int_{-\infty}^{\infty}da\int_{0}^{\infty}d_4k\,\frac{1}{2}\frac{a}{\lvert a\rvert}e^{i[(k^2+K^2)a+kz+\omega_0 z]}$
Chapter IV	No errata	
Chapter V		
4	16	should denote the energy levels by $E_{n_1}\cdots n_N\cdots$
13	17	is the non-relativistic limit of the exact equation (3.176")
19	8	... in principle be solved to express k in terms
	11	This expression for k may then ...
21	2, 3	$y = r\sin\theta\sin\varphi$ $= r\sin\theta\{\sin\omega_m\sqrt{1-\cot^2\alpha\cot^2\theta} + \cos\omega_m\cot\alpha\cot\theta\}$
22	28	Confining our attention to the transition...
40	4	The lack of dependence of (5.146)....
42	9	angular frequency ω_{nm} must be proportional...
49	14	$S^{rad} = \frac{1}{4\pi c^3}\frac{(n\times\ddot{\eta}^1)^2}{r^2}$ (5.181)

Chapter VI :

8	9	.... consists simply of a single charged....

Chapter VII.

1	12	$\|\rangle$. We have already.....
2	9	..... vectors $\|A\rangle$ and $\|B\rangle$ in S there exists..
	22	7) is also equal to b (a $\|A\rangle$). Since....
4	26	ping $S \to \mathcal{H}$ of the state vector....
8	15	$O^* = O,$ (7.25)
9	7	... When the equality holds we have $\Theta = 0,$
12		In lines 16, 19, replace $\mathcal{S}$ by S
18	7	... are also real. More generally, if
20		In lines 12, 14, 16 replace "discord" by "discard"...
22		In lines 10, 11 replace "orthogonal" by "brthonormal"
23	16	$= f(\xi')\langle \xi'', \delta'' \| \xi', \delta' \rangle^* = f(\xi')\langle \xi', \delta' \| \xi'', \delta'' \rangle$

24	3	$= \delta_{\xi'\xi''}\lvert\xi'',\gamma''\rangle = \delta_{\xi''\xi'}\lvert\xi'',\gamma''\rangle = \delta_{\xi'\xi}\lvert\xi'',\gamma''\rangle$
	5	$\delta_{\xi'\xi} \equiv \sum_{\gamma'} \lvert\xi',\gamma'\rangle\langle\xi',\gamma'\rvert$ (7.96)
	19	... The expansion is, of course, unique. An....
25	8, 9	$\xi(\eta-\eta')\lvert\xi',\gamma'\rangle = (\eta-\eta')\xi\lvert\xi',\gamma'\rangle$ $= (\eta-\eta')\xi'\lvert\xi',\gamma'\rangle = \xi'(\eta-\eta')\lvert\xi',\gamma'\rangle$
32	12	Omit the last word "have".
33	11	Change "have been" to "will be".
35	3	$\langle\lvert\delta_{\alpha'\alpha}\rvert\rangle = \sum_{\gamma'}\langle\lvert\alpha'\gamma'\rangle\langle\alpha'\gamma'\rvert\rangle$

Chapter X.-

Page

5 The limits of integration in the last term of Eq. (10.24) should be changed to read

$$\int_{t'}^{t''} dt_2 \int_{t'}^{t_2} dt_1$$

9 line 18. Add " α_i " after the word "operators"

14 Replace line 21 by "... often does not. We shall, in fact, find that $U_I(t, -\infty)$ frequently possesses matrix elements which diverge in the limit $\epsilon \to +0$. Let us suppose, for the time being, however, that $U_I(t, -\infty)$ remains ... "

27 Omit lines 3 to 9. The reasoning here is fallacious, since $\langle \alpha'_0 | H_{IS} | \rangle_{\alpha'_0} = \langle \alpha'_0 | H_{IS} U(0, -\infty) | \alpha'_0 \rangle = \frac{i}{2\pi} \langle \alpha'_0 | X_- | \alpha'_0 \rangle$. The error lies in incorrectly integrating by parts and ignoring "surface terms at infinity.

64 The proof beginning at the bottom of the page is actually trivial. The result may be obtained by writing

$$U_I^{(1)}(t_0, t) = e^{\frac{1}{i\hbar}(H_0 + H_1)(t_0 - t)} e^{\frac{1}{i\hbar} H_0 (t - t_0)}$$

$$U_I^{(2)}(t_0, t) = e^{\frac{1}{i\hbar} H (t_0 - t)} e^{\frac{1}{i\hbar}(H_0 + H_1)(t - t_0)}$$

$$U_I(t_0, t) = e^{\frac{1}{i\hbar} H (t_0 - t)} e^{\frac{1}{i\hbar} H_0 (t - t_0)}$$

$$= U_I^{(2)}(t_0, t)\, U_I^{(1)}(t_0, t)$$

and taking the limit $t \longrightarrow -\infty$. Feynman's "disentangling" techniques have thus been unnecessarily applied.

Page 73 line 4 Change "dumping" to "damping".

APPENDIX

The appendix should be retitled "The equivalence of Heisenberg's S-matrix with the Dyson transformation function $U_I(\infty, -\infty)$".

Page 82 (Footnote)

line 4 of the footnote : read eqs. (3.51) and (3.65).